National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

Ribosomal RNAs as Indicators of Relatedness Among *Pseudomonas* spp.

by

Neeraja Sankaran

©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF MICROBIOLOGY.

EDMONTON, ALBERTA

FALL, 1990

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-64971-2

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: **NEERAJA SANKARAN**

TITLE OF THESIS: **RIBOSOMAL RNAS AS INDICATORS OF RELATEDNESS AMONG *PSEUDOMONAS* SPP**

DEGREE: **MASTER OF SCIENCE**

YEAR THIS DEGREE GRANTED: **1990**

PERMISSION IS GRANTED TO THE UNIVERSITY OF ALBERTA LIBRARY TO REPRODUCE SINGLE COPIES OF THIS THESIS AND TO LEND OR SELL SUCH COPIES FOR PRIVATE, SCHOLARLY OR SCIENTIFIC RESEARCH PURPOSES ONLY.

THE AUTHOR RESERVES OTHER PUBLICATION RIGHTS, AND NEITHER THE THESIS NOR EXTENSIVE EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT THE AUTHOR'S WRITTEN PERMISSION.

(Student's Signature)
G-18, Sector 14,
Chandigarh, 160014.
India.

Date: 14 JAN 1990

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled RIBOSOMAL RNAs AS INDICATORS OF RELATEDNESS AMONG *PSEUDOMONAS* spp. submitted by NEERAJA SANKARAN in partial fulfillment of the requirements for the degree of Master of Science.

Dr. J.N. Campbell - Supervisor

Dr. K.L. Roy

Dr. S.E. Jensen

Dr. D. Cass

Date: August 15, 1990

ABSTRACT

The relationships among various closely related *Pseudomonas* species were studied by comparing partial sequences of their 23S ribosomal RNAs. Sequences were determined by a reverse transcriptase-catalyzed primer extension method. Synthetic oligonucleotides from relatively conserved regions of the 23S molecule were used as primers and rRNA was used directly as a template. Sequences of length ranging from 200-250 nt were obtained from each primer yielding a total of 1220-1400 nt per organism.

The identification of a hitherto unclassified organism *Pseudomonas* rp2 was possible by this method. Relationships among other closely related *Pseudomonas* spp., as well as more distantly related eubacterial species were also clarified. The RFLP technique, using labelled rRNA as a probe was also used in studying interrelationships among the various *Pseudomonas* spp. The results correlated well with the sequence data.

ACKNOWLEDGEMENTS

I would like to extend my heartful gratitude to Dr.Campbell and Dr. Roy for all their help and guidance throughout my course of my stay here. I would also like to thank the other members of my committee, Dr Cass and Dr. Jensen, for their helpful comments regarding my work.

I thank all of my friends and colleagues at the Dept. of Microbiology for a very enjoyable time here at Edmonton. Special thanks are due to Jeff Hoyem for all the times he got me out of jams with the computer, and to Domenic Spadafora and Sanja Saftic for "babysitting" me while I wrote my thesis. Also, I would like to thank my parents, brother and friends back home for all their long distance encouragement, and for believing in me.

TABLE OF CONTENTS

Abstract iv

Acknowledgements v

List of tables viii

List of figures ix

List of abbreviations xii

I. INTRODUCTION

Bacterial classification 1

Ribosomal RNAs as phylogenetic markers 4

Restriction fragment length polymorphisms 10

The classification of the genus *Pseudomonas* 14

Research objectives 17

II. MATERIALS AND METHODS

Bacterial strains and media 18

DNA and RNA purification 18

Restriction fragment length polymorphisms

(i) Digestion with restriction enzymes 20

(ii) Southern transfer 21

(iii) RNA probes 21

(iv) Hybridization 22

RNA sequencing

(i) rRNA templates 22

(ii) Oligonucleotide primers 22

(iii) Enzyme and nucleotides 23
(iv) Annealing reactions 23
(v) Separation of fragments. 25
Data manipulation 25

RESULTS
Restriction fragment length analysis 27
rRNA hybridizations 27
23S rRNA sequencing 40

DISCUSSION
RFLP analysis 89
Technical aspects of the sequencing analyses 90
Sequence alignments 93
Bacterial phylogeny 110
Conclusions 112

REFERENCES 113

LIST OF TABLES

Table 1.	Bacterial strains	19
Table 2.	Oligonucleotide primers	24
Table 3.	Pairwise comparison of the authentic *Pseudomonas* species.	65
Table 4.	Pairwise comparison of selected eubacterial species.	84

LIST OF FIGURES

Figure 1. Agarose gel electrophoretic separation of *Eco*RI digests of genomic DNAs. 28

Figure 2. Agarose gel electrophoretic separation of *Eco*RV digests of genomic DNAs. 30

Figure 3. Agarose gel electrophoretic separation of *Hin*dIII digests of genomic DNAs. 32

Figure 4. Autoradiograph of the genomic DNAs digested with *Eco*RI and probed with *Pseudomonas* rp2 rRNA. 34

Figure 5. Autoradiograph of the genomic DNAs digested with *Eco*RV and probed with *Pseudomonas* rp2 rRNA. 36

Figure 6. Autoradiograph of the genomic DNAs digested with *Hin*dIII and probed with *Pseudomonas* rp2 rRNA. 38

Figure 7. Secondary structure of the complete 23S rRNA molecule of *P. aeruginosa*.

7a. 5'-half 42

7b. 3'-half 43

Figure 8. Autoradiograph of a representative reverse transcriptase sequencing gel. 44

Figure 9. Partial sequence alignments of the authentic *Pseudomonas* species with primer #1. 46

Figure 10. Partial sequence alignment of the authentic *Pseudomonas* species with primer #6. 49

Figure 11. Partial sequence alignment of the authentic *Pseudomonas* species with primer #8. 52

Figure 12. Partial sequence alignment of the authentic *Pseudomonas* species with primer #3. 55

Figure 13. Partial sequence alignment of the authentic *Pseudomonas* species with primer #7. 58

Figure 14. Partial sequence alignment of the authentic *Pseudomonas* species with primer #5. 61

Figure 15. Partial sequence alignments of selected eubacterial species with primer #1 66

Figure 16. Partial sequence alignments of selected eubacterial species with primer #6. 71

Figure 17. Partial sequence alignments of selected eubacterial species with primer #3. 74

Figure 18. Partial sequence alignments of selected eubacterial species with primer #7. 78

Figure 19. Partial sequence alignments of selected eubacterial species with primer #5. 81

Figure 20. Unrooted phylogenetic tree showing the relationships among the authentic *Pseudomonas* species. 85

Figure 21. Unrooted phylogenetic tree showing the relationships among selected eubacterial species. 87

Figure 22. Secondary structure model of domain I of the 23S rRNA of *P. aeruginosa* 95

Figure 23. Secondary structure model of domain II of the 23S rRNA of *P. aeruginosa* 97

Figure 24. Secondary structure model of domain III of the 23S rRNA of *P. aeruginosa* 100

Figure 25. Secondary structure model of domain IV of the 23S rRNA of *P. aeruginosa* 102

Figure 26. Secondary structure model of domain V of the 23S rRNA of *P. aeruginosa* 104

Figure 27. Secondary structure model of domain VI of the 23S rRNA of *P. aeruginosa* 107

LIST OF ABBREVIATIONS

ddNTP	Dideoxyribonucleoside triphosphate
dNTP	Deoxyribonucleoside triphosphate
DTT	Dithiothreitol
EDTA	Ethylene-diaminetetra-acetic acid
min	Minutes
nt	Nucleotides
RNase	Ribonuclease
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SSPE	Sodium chloride/ sodium phosphate/ EDTA buffer[1]
TE buffer	Tris-HCl/EDTA buffer[2]
TEA buffer	Tris/EDTA/Acetate buffer[3]
TEB buffer	Tris/EDTA/Borate buffer[4]
Tris-HCl	Tris-(hydroxymethyl) aminomethane hydrochloride

1 The compositon of 20 x SSPE buffer is 3 M NaCl, 0.2 M $NaH_2PO_4.H_2O$, 0.02 M EDTA, (adjusted to pH 7.4 with NaOH).

2 The composition of TE buffer is 10 mM Tris-HCland 1 mM EDTA (pH 7.8).

3 The composition of 1 x TEA is 20 mM Tris acetate, 50 mM sodium acetate and 2 mM EDTA (pH 7.4).

4 The composition of 1 x TEB is 100 mM Tris, 100 mM boric acid and 1 mM EDTA (pH 8.3).

INTRODUCTION

Bacterial Classification

Bacterial classification is the process of grouping various microorganisms according to their similarities and differences. It is an important tool in scientific research, since it allows the organization of a large and diverse population into an orderly and manageable framework. In addition, it facilitates identification, which allows the exchange of information and avoids unnecessary repetitions of experiments. A number of criteria must be considered when arranging microorganisms into different groups, and the history of classification will be discussed briefly before dealing with the principles underlying modern taxonomic procedures.

Originally, classification systems were based entirely on phenotypic characteristics of the microorganisms. Initially methods relied heavily on morphology, and as the importance of physiological features became apparent they too became important tools in classification. However, classification was still an empirical process where organisms were characterized as thoroughly as possible in terms of their morphological, physiological, and biochemical properties, (cell shape, size, nutritional modes, oxygen requirements, and sugar fermentations are some representative examples) and then arranged into groups based on the intuitive judgement of the individual scientist. Many of these characterizations have proven valid until the present; the most important being the Gram staining character of the bacterial cells. This stain, which distinguishes the bacteria on the basis of the peptidoglycan content of their cell walls, is the basis for dividing all the eubacteria into two groups: gram positive and gram negative. In general, however, this classification system had a major disadvantage in that it was very subjective. Different scientists gave varying weights to

different characters and this often led to conflicting classification schemes. Thus, a more objective and standardized system of classification was required.

Objectivity was achieved by numerical methods, where each characteristic was given equal weight. The data took the form of tables which compared the various physiological and biochemical properties of the different organisms. Clustering patterns among the organisms could be determined on analyzing these data. The problem of handling such data was solved by using computers, which could analyze extensive tables more efficiently than could be done by hand. Various programs have been developed for this purpose, e.g. TAXMAP (used by Campbell *et al.* 1986). The clusters or taxa defined by numerical taxonomy were polythetic in nature in that members of a group shared a high proportion of attributes but not necessarily any invariable set. In addition to providing a means of classifying known bacteria, this scheme allowed the placement of new isolates within the existing framework, based on a comparison of their properties.

The earlier methods of classification were phenetic in nature. This means that the organisms were divided into groups according to their existing similarities and differences. The source of these attributes was not taken into account. Natural or phylogenetic relationships, on the other hand, are defined by their evolution. The origin of the observed properties is an important consideration. Differences are observed in the genomic sequences of the organisms and are expressed in terms of changes in an evolutionary lineage. Relationships may be deduced from existing similarities and differences by assuming that evolution has proceeded in an orderly fashion. There are some disadvantages to using this approach since two organisms can share several attributes even though they descend from unrelated lineages. Also various phenotypic properties are subject to environmental influences, which can lead to erroneous interpretations. Natural relationships are better determined by an examination of the genotypes of the organisms under consideration.

All classification schemes use a system of ranking to categorize the various members of the group. Bacterial classification resembles that of higher plants or animals in its structure. According to *Bergey's Manual of Systematic Bacteriology* (1984), all prokaryotic organisms are placed in the Kingdom *Procaryotae*, and are further subdivided into various groups according to the following hierarchy: Division, Class, Order, Family, Genus, and Species. An organism is identified by a binomial name consisting of its genus and species. A strain refers to an individual isolate and is derived from a single colony from a pure culture. Thus, a species consists of several individual strains which may or may not be identical to one another. A single species may be further subdivided into subspecies according to minor but consistent variations occurring within the group. Certain special characters may sometimes be used to distinguish among a group of very closely related strains; common examples are serovars which are defined on the basis of their antigenic makeup, or pathovars which are so designated because of their differential pathogenicity towards a range of hosts. While such rankings serve as practical tools in diagnosis and epidemiological studies, they are often purely empirical and may show little correlation with the phylogenetic relationships among bacteria.

There are several approaches to comparing the genotypes of bacteria. Earlier methods involved comparisons based on DNA composition or DNA-DNA hybridizations. The DNA composition of an organism is calculated in terms of its constituent bases by the mathematical formula. $[(G + C) \div (G + C + A + T)] \times 100$. It is generally referred to as the %G+C value and is constant for a given organism. Different bacteria fall within a wide range of %G+C values (25-75 %) and similar values may imply relatedness. The members of a given taxonomic unit are assumed to fall within a narrow range (*Shigella* species fall within a range of 49-53 mol%) but several exceptions exist. An example is the genus *Spirochæta* which has a wide range of %G+C values from 51-65 mol%. Two bacteria can have the same %G+C with

their genes specifying totally unrelated functions. *Micrococcus* and *Streptomyces* are unrelated genera with similar %G+C values in the range of 64-75%. It is the similarity in the sequence of the bases rather than their total content that indicates the extent of relatedness among various organisms.

DNA-DNA hybridization is an indirect way to compare the sequences of two organisms. When double stranded DNA molecules are subjected to high heat or pH, the complementary strands dissociate from each other (denature). The temperature at which 50% of all the strands separate is termed as its melting point (T_m). It is specific for a given DNA molecule and depends mainly on its G+C content. Under conditions of lower temperature and high salt concentration, the single strands reanneal to one another. If denatured genomic DNAs from two different organisms are mixed and allowed to renature they will anneal with each other in varying extents to form heteroduplexes. The thermal stability of the hybrid thus formed is measured by determining the T_m. The lower the homology between two species, the greater the number of mismatched base pairs in the hybrid, and hence the lower the melting point (Palleroni, 1983). This method gives information on the relative sequence similarities of the two species but the data obtained are not as precise as actual sequences. Ideally, one should compare entire sequences for the accurate identification and classification of any organism, but this is not at all a practical approach due to the time required and the overwhelming amount of data involved. Therefore, only specific genes or regions of the genome can reasonably be used to derive phylogenetic relationships among bacteria.

Ribosomal RNAs as phylogenetic markers

The genes chosen for sequence analyses have to meet certain criteria for the information to be of any taxonomic significance; they should be present in each of the

individuals being examined and the gene products should all have similar functions. Because of their universal presence in living entities and constant function as the protein synthetic machinery in all cells, the ribosomes have been targeted as the most useful structure for deriving phylogenetic relationships. Their constituent molecules are known to change rather slowly over time (Fox *et al.* 1980). Thus, the greater the similarity between the ribosomes of two organisms, the less the evolutionary distance between them.

Ribosomes are complex structures consisting of both protein and RNA components. Prokaryotic and eukaryotic ribosomes show essentially the same features, the major difference being in their size (80S in eukaryotes and 70S in prokaryotes). The prokaryotic ribosome is composed of a large (50S) and a small (30S) subunit. There are 3 distinct RNA species distinguishable by their sedimentation coefficients: the 23S and 5S species in the large subunit and the 16S RNA in the small subunit. The ribosomal RNAs (rRNAs) are associated with the ribosomal proteins in a complex manner. rRNA genes are usually present in multiple copies per genome, an occurrence otherwise uncommon in bacteria. All three rRNA species are normally encoded in the same operon to ensure equimolar quantities in the cell at all times.

RNA molecules are preferred for comparison studies because they are with some exceptions, a direct reflection of the DNA sequences, unlike protein molecules whose sequences are open to erroneous interpretations due to codon degeneracies (Hori, 1975). The use of rRNA homologies in exploring bacterial phylogenies has been well established. Various studies have been conducted using different rRNA components and techniques, depending on the nature of the group of organisms under scrutiny. Some of the methods that have been used in the past are (a) rRNA-DNA hybridizations, (b) oligonucleotide cataloging of 16S rRNA (Fox *et al.* 1980), and (c) total sequence comparisons. Johnson and Harich (1983) compared the clustering

patterns of some *Clostridium* species using procedures (a) and (b) and found the same outcome in both cases.

RNA-DNA hybridizations have been used extensively in the classification of various groups of bacteria. (DeLey *et al.*, 1978; DeSmedt and DeLey, 1977; DeSmedt *et al.*, 1980; DeVos and DeLey, 1983; DeVos *et al.*,1985; DeVos *et al.*, 1989; Jarvis *et al.*, 1986). Labeled rRNA from one organism (the reference strain) is hybridized to fixed amounts of immobilized DNA from a number of species. Unbound RNA is removed by S1 nuclease, the amount of bound RNA is estimated by measuring the radioactivity, and then thermal denaturation curves are plotted for each DNA-RNA hybrid. Two parameters are used to characterize the hybrids; (a) $T_{m(e)}$ value, which is the temperature at which 50% of the bound RNA is eluted, and (b) the extent of hybridization, expressed as % rRNA binding, (µg rRNA per 100 µg DNA). These parameters are then used as the two axes of an RNA similarity map, where members of a taxonomic unit form clusters (DeVos and DeLey, 1983). The main property for this characterization is the $T_{m(e)}$ value, since it is directly affected by the number of mismatched bases between the ribosomal gene and the reference rRNA. The %rRNA binding however, depends on the number of rRNA cistrons in the genome rather than on the relative homologies and hence is not as reliable an indicator of relatedness. This method has proved useful in elucidating relationships among various bacteria at the generic and suprageneric level, but has severe limitations when applied to closely related species, because a small difference in homology is not accurately reflected by a corresponding change in the $T_{m(e)}$. DeVos and DeLey (1983) have reported that a difference of 1°C in $T_{m(e)}$ from ribosomal DNA similarity was equivalent to roughly 14% DNA homology. Also, as in the case with DNA-DNA hybridizations, no sequence data are obtained and hence the precise locations of the mismatched bases remain unknown.

Oligonucleotide cataloging involves the characterization of the RNA in terms of sequences of short oligonucleotides. This method was the original approach to comparing sequences directly. The rRNAs were digested with T1 ribonuclease (which cleaves specifically on the 3' side of every guanylate residue) to yield a mixture of oligonucleotides. Fragments of 5 nucleotides or longer were resolved by 2-dimensional electrophoresis and sequenced, thus producing a catalogue of oligonucleotides specific for a given organism. These catalogues were then compared to derive evolutionary relationships among various bacteria. The data were analyzed in terms of binary association coefficients (S_{AB} values), calculated for each pair of bacteria (Fox *et al.*, 1980). S_{AB} is calculated by the mathematical formula: $2N_{AB} + (N_A + N_B)$, where N_A and N_B are the total number of bases in all the oligonucleotides (of 5 bases or longer) in catalogues A and B respectively, and N_{AB} is the total number of bases in the oligonucleotides common to both catalogues (Woese 1987). Dendrograms were then derived from these values by cluster analysis (Fox *et al.*, 1980; Woese 1987). Organisms were grouped on the basis of their similarities to each other and the average linkage between merged groups was computed.

With the advent of rapid sequencing techniques oligonucleotide cataloging has become obsolete. Sequences are determined from either the cloned rRNA genes (Toschka *et al.*, 1988) or the ribosomal RNAs themselves (Lane *et al.*, 1985), and aligned for comparison purposes. Percentage homologies are determined, based on the number of mismatches in the alignable portions of the RNA, and evolutionary distances (K_{nuc}) are calculated between the individuals. $K_{nuc} = -3/4 \ln \{ 1 - (4/3) \lambda \}$, where λ is the fraction of mismatches between the sequences of two organisms (Hori, 1975). The algorithm provided by De Soete (1983), can then be used to construct dendrograms from evolutionary distance matrices (Bowen *et al.*, 1989; Weisberg *et al.*, 1989). An alternative method for the construction of evolutionary trees is derived from difference matrices between the isolates (Klotz and Blanken, 1981). This method yields unrooted

trees showing the phylogenetic distances between present day species of bacteria. It has an advantage in that complex numerical operations are not required.

Comparative rRNA sequence analyses have been conducted on various rRNA components. The evolution of 5S RNA has been studied by Hori (1975), using sequences from both eukaryotic and prokaryotic sources. The overall pattern of divergence obtained correlated with previous studies based on cytochrome c analysis. 5S RNA sequence comparisons were also used by Stahl *et al.* (1984) as a rapid means of establishing the phylogenetic affiliations of certain bivalves and their bacterial symbionts from hydrothermal vents. Van den Eynde *et al.* (1990) have discussed the the evolution of the green sulphur bacteria on the basis of their 5S rRNA sequences. In general, however, larger components (16S and 23S components in prokaryotes, 18S and 25 or 28S in eukaryotes) are more useful in phylogenetic studies because they provide more statistically significant data; the typical size of a 5S rRNA molecule is 120 nucleotides, in contrast to 16S and 23S RNAs which are around 1.5 and 2.9 kb, respectively. Thus, a percent homology of 80% between the RNAs of 2 organisms is reflected as 24 individual differences in the 5S, 300 in the 16S, and 580 in the 23S rRNA molecule. Besides their greater information content, the larger rRNA molecules have a characteristic mosaic-like structure with adjacent domains evolving at independent rates, which makes them more useful in phylogenetic studies than the 5S component (Baroin *et al.* 1988).

The most widely used RNA component in these studies has been the small subunit RNA (16S in prokaryotes and 18S in eukaryotes). Lane *et al.* (1985) demonstrated that it was not necessary to compare entire RNA sequences. They sequenced specific regions of the molecule using a primer extension/dideoxynucleotide method. The enzyme reverse transcriptase was used for chain elongation with purified rRNA as template. Three "universal" sites, 15-20 bases in length were chosen on the 16/18S rRNA molecule, and oligonucleotides complementary to them were used as

primers for sequencing. These priming sites were chosen since they showed strong interkingdom conservation. The 16S sequences of *Escherichia coli*, *Halobacterium volcanii*, and *Dictyostelium discoideum* were used as representatives of the eubacteria, archaebacteria and eukaryotes respectively, to find these regions. The hybridizing locations on *E. coli* 16S RNA were 519-536, 907-926, and 1392-1406. The phylogenies derived from these partial sequences were in good agreement with those based on complete sequence comparisons.

This method of comparing partial rRNA sequences is gradually becoming a major taxonomic tool. The 16S component of rRNA has been used extensively in the classification and identification of microorganisms (Bowen *et al.*, 1989; Schleifer *et al.*, 1990; Weisberg *et al.*, 1989; Stackebrandt and Charfreitag, 1990; Zhao *et al.*, 1990). The large subunit rRNA has also been utilized in phylogenetic analyses of various organisms. e.g. 23S rRNA of archaebacteria (Leffers *et al.*, 1987), 25S rRNA of the basidiomycetous yeasts (Gueho *et al.* 1990), and the 28S rRNA of certain eukaryotic protists (Baroin *et al.*, 1988). Since the 23S-like molecules are larger and show a greater degree of variability than the 16S-like molecules, their sequences are more useful for inferring relationships among bacteria within a kingdom, family or even a genus. By the same rationale, 16S-like sequences would probably yield more accurate phylogenetic trees between kingdoms (Leffers *et al.* 1987). The overall phylogenetic tree for eubacteria and archaebacteria based on 23S sequence data (Höpfl *et al.*, 1989) is similar to that obtained by 16S sequence analyses (Woese, 1987).

As mentioned earlier, the 16S and 23S rRNAs have a mosaic-like structure composed of regions of varying sequence conservation. Such a pattern is seen because of the extensive secondary structure of ribosomal RNAs. The first secondary structure model for the large subunit RNAs was proposed by Noller *et al.* (1981) using the primary sequence from *E. coli*. Based on comparisons of 23S and 23S-like rRNAs from archaebacteria, eubacteria, eukaryotic organelles and cytoplasms, this basic model

has been refined. Gutell and Fox (1988) have presented several of these sequences in a structural format, based on this secondary structure model. The 23S molecule consists of six loosely associated structural domains (numbered I to VI from the 5'-end), each with several helices and loops. These domains show several characteristic features. For example specific nucleotide positions have been identified in domain V of archaebacterial rRNA, which are associated with certain antibiotic binding sites in the ribosome (Leffers *et al.*, 1987). Seven tertiary structural interactions, common to all 23S-like rRNAs were predicted within unpaired regions of the secondary structure model (Leffers *et al.*, 1987), and phylogenetic evidence for six of these has since been provided (Höpfl *et al.*, 1989). A detailed analysis of the conservation and variability of each domain of the 23S molecule has been conducted, based on the sequence data from the compilation by Gutell and Fox (1988) and additional eubacterial and archaebacterial sequences. In general the 5' half of the 23S RNA has more variable regions than the 3' half (Höpfl *et al.*, 1989). Different domains are useful in different situations; faster evolving regions are of value when studying a closely related population, while domains with a slower rate of change are more advantageous when examining relationships among distant organisms.

Restriction fragment length polymorphisms

While ribosomal RNA sequences are very reliable and convenient indicators of relatedness, they are by no means the sole source of evolutionary information. Another very useful technique is DNA fingerprinting. This method takes advantage of the susceptibility of any DNA molecule to restriction endonucleases. These enzymes recognize specific short palindromic sequences (usually 4-6 bases long) and cleave the

DNA molecule at these sites. Thus, upon digestion with a given restriction endonuclease a DNA molecule may be cut into a number of fragments of variable length. These fragments can then be separated on agarose or polyacrylamide gels according to their size and visualized by ethidium bromide staining. Differences in the length of a particular fragment may arise due to a number of possible genotypic changes. It may be altered in size due to an insertion or deletion within its sequence. One or more individual bases might differ between two organisms, resulting in the loss of a cleavage site or the creation of a new one. Such genotypic changes are recognized by the altered mobility of the restriction fragments on a gel (Botstein *et al.*, 1980). The genomic differences revealed by this technique are termed restriction fragment length polymorphisms (RFLPs). They serve as important genetic markers, and have been used in various types of genetic analyses, including genetic mapping (Botstein *et al.*, 1980), and in taxonomy (Bautsch *et al.* 1988; Cook and Sequeira, 1988; Denny *et al.* 1988; Gottleib and Rudner, 1985, Lazo *et al.* 1987; Levy-Frebault *et al.* 1989; Poh *et al.* 1989).

DNA fingerprints are useful in distinguishing among various organisms since each organism exhibits a characteristic profile of fragments or bands, depending on the location and distribution of the recognition sites of the restriction endonuclease being used. Closely related organisms would generally have comparable digestion profiles since the location of the susceptible sites on the genome would be similar. The endonucleases used for RFLP analyses should be chosen such that the number of fragments generated are within a range that can be resolved on a gel. The enzymes *Bgl*II and *Hin*fI were used to characterize various strains of *Neisseria gonorrhoeae*, (Poh *et al.*, 1989). *Hin*fI yielded fragments that were very well resolved in the size range of 2 to 6.5 kilobases, while *Bgl*II showed differences in patterns over a wider range of fragment sizes (2 to 21.5 kb). The strains could be differentiated into groups on the basis of the presence and absence of specific bands. The endonuclease *Dra*I was

similarly utilized to compare various *Mycobacterium* species (Levy-Frebault *et al.*, 1989). Infrequent cutters yield only a few fragments which are usually too large to resolve by simple agarose gel electrophoresis and may require special techniques like pulsed-field gel electrophoresis. Also, polymorphisms may not be detectable in these few sites. Very frequent cutters may produce too many fragments, making the interpretation of the results difficult. One way to solve the latter problem is to visualize only a small fraction of the total DNA. This can be achieved by transferring the digested DNA onto nitrocellulose or nylon membranes and then hybridizing with a specific radioactive probe. The probe will bind to complementary regions in the genome, which are then detected by autoradiography. This modification of the basic RFLP technique has been used with success in many cases. Lazo *et al.* (1987) used it to distinguish various pathovars of *Xanthomonas campestris*. Gottleib and Rudner (1985) used this technique to differentiate among species of the genus *Bacillus*.

Genetic variation can be observed for random or selected target sequences within a species. The choice of the probe sequence is based on the population being tested. For reasons stated earlier (constant occurrence and function), ribosomal RNA sequences are very useful as probes in RFLP analyses. Arthur *et al.* (1990) compared the RFLP patterns among various uropathogenic isolates of *E. coli*, using probes from the virulence associated genes (*pap* -related sequences) and the ribosomal RNA (*rrn*) operons. While the variations in the *rrn* RFLPs were consistent with previously defined evolutionary groups, the *pap*-related RFLPs showed considerably more divergence. An important conclusion that could be drawn from these data was that the *pap* -related sequences evolved at a much more rapid rate than did the rRNA genes. Isolates with similar *pap* patterns invariably belonged to the same electrophoretic group defined by the *rrn* probes. Lazo *et al.* (1987) selected random clones from a genomic library of *X. campestris* pv. *citri* to probe different strains of the organism. While some clones represented genetic loci that were strongly conserved within the species,

some others represented highly variable loci. The method was also used to identify *Campylobacter* spp using a conserved sequence from 16S rRNA as a probe (Moureau, *et al.*, 1989). This method is most useful in distinguishing among very closely related strains of bacteria and in confirming the identity of an organism whose taxonomic affiliation is suspected. It cannot be used to screen very heterogenous populations, and since the data are essentially qualitative in nature they do not lend themselves well to numerical or statistical analyses.

RFLPs may be exploited to compare the entire sequences of two organisms without having to sequence either genome. Such a comparison is achieved by cross hybridization (Shen and Maniatis, 1980; Keen *et al.*,1988). DNA samples from two organisms are digested with an enzyme, and one sample is electrophoresed on an agarose gel, then immobilized on nitrocellulose or nylon. The second DNA sample is radioactively end-labelled prior to separation on a gel, after which it is hybridized with the first sample at a 90° orientation. If the DNAs are identical, autoradiography of the membrane would reveal a diagonal pattern of spots, with a few extra spots placed symmetrically about the diagonal indicating repeated sequences. Closely related species should yield a greater number of spots along the diagonal, while more distant organisms would display a greater degree of dissymmetry. If the two organisms have very little homology then a great number of blank areas would be observed due to a lack of hybridization. The method has been used to study specific cloned genes like the rabbit β-like globin sequences (Shen and Maniatis, 1980), and small DNAs like plasmids (Keen *et al.*, 1988).

While this procedure has potential as a means of rapidly comparing two organisms and detecting small differences between them, it poses certain technical problems which would prevent its use as a widespread tool in classification. Firstly, the use of this technique is very labor intensive and time consuming, especially when handling large numbers of organisms. Each comparison requires 2 gels and takes at

least two days for completion (2 sets of transfer reactions are required per pairwise comparison). Each new isolate would entail a pairwise comparison with all pre-existing organisms and this would quickly generate unmanageable amounts of data. In addition, an extremely high specific activity labelling of the probe DNA is needed to visualize the hybridization spots. This level of labelling is not easily obtained when complete genomic DNAs are used.

The classification of the genus *Pseudomonas*

For over three decades, the classification of the members of the genus *Pseudomonas* has been a subject of interest to microbial taxonomists. Typically these organisms are gram-negative, aerobic, non spore forming, rod shaped bacteria with polar flagella, and have DNA with a G+C content in the range of 58-69 moles%. They are ubiquitous in nature and are either chemoorganotrophs or facultative autotrophs that utilize molecular hydrogen as an energy source (Palleroni, 1983). This description is broad enough to include a large number of organisms and consequently the genus *Pseudomonas* consists of a heterogeneous group of species from a variety of habitats. Over the years the subject of *Pseudomonas* classification has been reviewed exhaustively and so the following discussion will deal with only a few of the representative papers.

Sands *et al.* (1970) conducted a detailed study on the pseudomonads in an attempt to organize them into coherent groups and facilitate proper identification, particularly of the phytopathogenic organisms. All the isolates were placed into four major groups based on their ability to fluoresce, their growth rates, the production of the enzymes oxidase and arginine dihydrolase, and the induction of hypersensitivity in tobacco plants. This last test was used to gauge the phytopathogenicity of the bacteria and a positive reaction was indicated by the death of injected plant tissue within 24

hours. A battery of over 50 nutritional and physiological tests were used in order to further characterize the relationships within these groups. Subsequent studies dealing with DNA homologies among *Pseudomonas* species (Palleroni *et al.*, 1972), and rRNA-DNA homologies (Palleroni, 1983; DeVos *et al.*, 1985), largely confirmed the conclusions of these phenotypic analyses. According to the scheme in *Bergey's Manual of Systematic Bacteriology* (1984), the family *Pseudomonadaceae* is composed of 4 genera: *Pseudomonas, Xanthomonas, Zoogloea,* and *Frateuria*. The genus *Pseudomonas* is divided into 3 Sections (I, II, and III) which are quite distinct from one another. The three sections are represented by the species *P. aeruginosa, P. solanacearum,* and *P. acidovorans* respectively.

Based on rRNA-DNA hybridization data, DeLey *et al.* (1978), defined 5 rRNA superfamilies amongst the gram negative bacteria. As a rule, all of the species of a genus cluster within a single superfamily. The notable exception in this case is *Pseudomonas*, where the different sections fall under different superfamilies. Organisms of Section I are seen to cluster with *Xanthomonas, Azotobacter* and *Azomonas* in the second superfamily, and have been termed as the authentic *Pseudomonas* species. Those grouped in Section II (the *P. solanacearum* branch) and Section III (the *P. acidovorans* branch) are members of the third superfamily along with *Alcaligenes* and *Janthinobacterium*, and are as different from each other as they are from the other two genera. Accordingly, DeVos and DeLey (1983) have proposed that organisms in Sections II and III are not *Pseudomonas* and should be renamed. The rRNA hybridization work also clarified the taxonomic positions of certain other isolates; *P. maltophilia* was shown to be a *Xanthomonas* (*X. maltophilia*) and *P. diminuta* and *P. vesicularis* were wrongly identified as *Pseudomonas* species.

Subsequent sequence analyses have further strengthened the hypothesis that the genus *Pseudomonas* consists of a heterogenous group of species. *E. coli* , *P. aeruginosa* and *P. testosteroni* (Section III), show a similar degree of 16S homology

to one another (~82-85 %), while *E. coli* shares 95% homology with *Proteus*, a different genus (sequence data from Dams *et al.*, 1988). Similarly, the 23S rRNA sequence of the Section II organism, *P. cepacia* (Höpfl *et al.*, 1989) is equally distant from both *E. coli* and *P. aeruginosa* (Gutell and Fox, 1988). As stated earlier, the distances between closely related organisms (within a genus) are not resolvable by rRNA-DNA hybridizations alone. However, rRNA sequence analysis should prove more discriminatory in that respect.

RFLP analyses have also been conducted on *Pseudomonas* species, especially while studying interrelationships at the genus or species level. Cook *et al.* (1989) classified pathovars of *P. solanacearum*, using DNA probes specifying virulence and the hypersensitive response. The genetic diversity and relationships among pathovars of *P. syringae* were examined by Denny *et al.* (1988). They digested DNA samples from different strains of 2 pathovars of *P. syringae* (*P. s. syringae* and *P. s. tomato*) and probed these with radiolabelled clones from a cosmid library of *P. s. tomato*. The experiments showed the two pathovars to be quite distinct. Highly conserved species-specific loci like the ones described by Lazo *et al.* (1987) in *X. campestris* pathovars were not detected in the strains of *P. syringae*. A higher degree of diversity was seen among strains of *P. s. syringae*, which have a wide host range, than among *P. s. tomato* strains, which have only one host. This work demonstrates the need for a standardized nomenclature of the phytopathogenic *Pseudomonas* species, based on more standardized classification methodologies. Many of the so called pathovars of *P. syringae* are often as distinct from one another as they are from other closely related species.

Research objectives

Originally, interest in this project was aroused while working with certain plant pathogenic pseudomonads. Campbell *et al.* (1986) isolated a soil microorganism from canola roots. It was an opportunistic pathogen capable of infecting only immature or sprouting roots. Preliminary tests revealed it to be a fluorescent *Pseudomonas* (Section I) but species identification was not possible. Temporarily designated as *Pseudomonas* rp2, the isolate showed no clustering with any of the known species, based on conventional biochemical and plant pathogenicity tests. More definitive procedures were therefore required to identify this isolate.

This thesis describes the work done in attempting to find the proper taxonomic location of *Pseudomonas* rp2. Two major lines of investigation were followed: 23S rRNA sequence comparisons and RFLP analyses. The methods are evaluated with respect to their general applicability in identifying and classifying new isolates.

MATERIALS AND METHODS

Bacterial strains and media

The bacterial strains used in this study are listed in Table 1. Cultures were grown on Trypticase Soy Agar plates (media from BBL) at 30°C for 24 hours. They were stored at 4°C and transferred at 2-3 week intervals. For DNA and RNA extractions the cells were grown in Trypticase Soy broth (BBL) at 30°C for 16-18 hours, with vigorous shaking to maintain aeration.

DNA and RNA purification.

Genomic DNAs and ribosomal RNAs were extracted from exponentially growing bacterial cells. The cells were centrifuged for 10 min. at 3020 x g. For DNA preparations pellets were resuspended in 25% sucrose, 50 mM EDTA (pH 8.0), (20 ml per gram of cells), treated with 0.5 mg/ml lysozyme for 15 min at room temperature followed by the addition of SDS to 0.5%, and incubation at 60°C for 30 min, and then 0.2 mg/ml proteinase K at 37°C for 3 hours. Proteins were removed by repeated extractions with phenol and DNA was precipitated from the aqueous phase by adding 2-3 volumes of ethanol. The precipitate was collected by centrifugation for 10 min at 7740 x g, dissolved in 10 mM Tris-HCl (pH 7.8) 1 mM EDTA,(TE buffer) and treated with 0.1 mg/ml of bovine pancreatic RNase (previously heated to 90°C for 10 min) at 37°C for 1 hr. The enzyme was removed by further phenol extractions and a final purification was achieved by dialysis against 0.2 M NaCl in TE buffer. The purified DNA samples were stored at -20°C.

TABLE 1. Bacterial strains.

Organism	Strain
Lysobacter enzymogenes	ATCC 29487
Pseudomonas acidovorans	ATCC 15668
Pseudomonas aureofaciens	Strain B-1534p
Pseudomonas cichorii	ATCC 10857
Pseudomonas fluorescens	ATCC 13525
Pseudomonas marginalis	ATCC 10844
Pseudomonas syringae pv. *maculicola*	ATCC 11781
Pseudomonas syringae pv. *syringae*	ATCC 19310
Pseudomonas syringae pv. *tabaci*	ATCC 17914
Pseudomonas viridiflava	ATCC 13223
Pseudomonas rp2	Field isolate*
Pseudomonas species (canker)	Field isolate*
Pseudomonas species (wheat)	Field isolate*

* Reported by Campbell *et al.* (1986).

RNAs were prepared according to the method described by Lane *et al.* (1985). Cell pellets were resuspended in 50 mM Tris-HCl (pH 7.8), 5 mM $MgCl_2$, 0.5 M NH_4Cl (10 ml buffer per gram of cells) with 20 µg/ml DNase. The cells were lysed in a French pressure cell and the lysate centrifuged at 48400 x g for 30 min at 4°C. The supernatant was extracted with phenol and RNA precipitated from the aqueous phase using 3 volumes of ethanol at -20°C. The RNA was collected by centrifugation (10 min, 7740 x g), and dissolved in TE buffer. High molecular weight ribosomal RNA was purified by precipitation in 2 M NaCl at 4°C for 16-18 hours. The precipitate was collected by centrifugation and redissolved in TE buffer. Samples were stored in ethanol at -80°C to minimize degradation.

The hot phenol /SDS method (Scherrer and Darnell, 1962) was used as a method for the rapid extraction of small quantities of RNA. Small aliquots (1.5 ml) of cultures in their exponential phase were pelleted in a microcentrifuge and each pellet resuspended in 50 µl of TE buffer + 0.3% SDS. An equal volume of phenol was added and the cells incubated at 70°C for 10 min, with vigorous mixing every 2-3 min. 100 µl of sterile distilled water was added to each tube and the aqueous and phenol phases separated by a brief centrifugation. DNA was spooled out from the aqueous phase after adding 2-3 volumes of ethanol. RNA was then precipitated from the resulting mixture by adding 1 volume of 5 M ammonium acetate and incubating at -20°C for 20 min.

Restriction Fragment Length Polymorphisms

Digestion with restriction enzymes

Restriction endonucleases (*Bam*HI, *Eco*RI or *Eco*RV) were purchased from Boehringer Mannheim Ltd. and were supplied with appropriate reaction buffers. Approximately 5 µg of each DNA sample was treated with 1-2 units of restriction

endonuclease per µg DNA for 2-3 hours at 37°C. The digested samples were fractionated by electrophoresis on 0.75% agarose gels (20 x 20 cm) in TEA buffer at 80 volts for 6-8 hours. The gels were stained in ethidium bromide (diluted to 0.1 µg/ml) and DNA was visualized under UV light. Lambda DNAs digested with either *Cla*I or *Bst*EII were used as a molecular weight marker to estimate fragment sizes. The gels were photographed using Polaroid type 665 film and an orange filter (Kodak Wratten gelatin filter number 23A).

Southern transfer

Nylon membranes (Hybond™N) were purchased from Amersham. DNA was transferred to the membranes after treatment of the gel as follows: depurination in 0.25 M HCl for 10 min, alkaline denaturation in 0.5 M NaOH, 1.5 M NaCl for 25 min, and neutralization and equilibration in 1.0 M ammonium acetate, 20 mM NaOH for 25 min (Rigaud *et al.*, 1987). The nylon membrane was prepared by wetting in 1.0 M ammonium acetate, 20 mM NaOH. The transfer was allowed to proceed in this solution for several hours after which the membrane was washed in 2 x SSPE for 10 min and dried under vacuum at 70°C for 30-60 min. (20 x SSPE: 3 M NaCl, 0.2 M $NaH_2PO_4.H_2O$, 0.02 M EDTA, pH 7.4 with NaOH).

RNA probes

RNA probes were prepared from the rRNA of *Pseudomonas* rp2. Approximately 100 µg of the RNA was precipitated, redissolved in 200 µl of 0.1 M $NaHCO_3$ / Na_2CO_3 (pH 9.0) and heated at 90°C for 15 min. 1.0 µl of this alkalline digested rRNA was end-labelled using 1-5 µl of [γ-^{32}P]ATP in 1 x kinase buffer (10 x buffer: 0.5 M Tris-HCl (pH 7.6), 0.1 M $MgCl_2$, 1.0 mM EDTA, 50 mM DTT, 1 mM spermidine) with 1-2 units T4 polynucleotide kinase for 30 min at 37°C. The

reaction was stopped by adding 200 µl 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.1% SDS.

Hybridization

Nylon membranes were prehybridized in 6 x SSPE, 5 x Denhardt's solution and 0.1% SDS at 60°C for 3-6 hours. Hybridization was carried out at 45-50°C for 16-18 hours in a 1:1 mixture of the above solution and formamide. Membranes were washed twice in 2 x SSPE + 0.1% SDS at 50°C for at least 15 min each, wrapped in Saran Wrap and autoradiographed at -80°C. 50 x Denhardt's solution is 10 g/l each of Ficoll (MW 40,000), polyvinyl pyrrollidone and bovine serum albumin (Denhardt, 1966).

RNA Sequencing

rRNA templates

rRNA templates were prepared as described previously. They were dissolved to a final concentration of approximately 2 mg/ml, for use in sequencing reactions.

Oligonucleotide primers

Primers for sequence analysis were chosen on the basis of comparisons made between the 23S rRNA sequences of *E. coli* and *P. aeruginosa* (Gutell and Fox, 1988). Conserved regions, 16-18 bases long, lying adjacent and upstream (towards the 3' end) of more variable regions were chosen as priming sites. Complementary oligonucleotides were synthesized by an Applied Biosystems Model 381A DNA Synthesizer. The primers, their sequences, and hybridization locations are listed in

Table 2. The primers were dissolved in distilled water and a small aliquot diluted to approximately 2 pmoles/µl for use in sequencing reactions. 1 µl of the diluted oligonucleotide was end-labelled using 3 µl of fresh [γ-^{32}P]ATP at 2 pmoles / µl in 1 x kinase buffer with 1-2 units of polynucleotide kinase.

Enzyme and nucleotides

AMV reverse transcriptase (RT) was purchased from Pharmacia or Boehringer Mannheim Ltd. and stored at -20°C. The enzyme was diluted to 2 units / µl in 1 x RT buffer (5 x RT buffer is 250 mM Tris-HCl (pH 8.3), 300 mM NaCl, 150 mM $MgCl_2$, 50 mM DTT) just prior to use. dNTPs were diluted to 2 mM each in 1 x annealing buffer (5 x annealing buffer is 250 mM Tris-HCl (pH 8.3), 300 mM NaCl, 50 mM DTT) and ddNTPs were diluted to 250 µM in 1 x annealing buffer for use in the sequencing reactions.

Annealing reactions.

An aliquot of 3-5 µl of labelled primer (corresponding to 1 pmole) was added to 3 µl of template (at 2 mg/ml) in 1x annealing buffer (final volume 10 µl). The mixture was heated at 90°C for 1 min and cooled to 50°C for 2-3 minutes, allowing the primer and template to anneal. While the annealing reaction was proceeding, the tubes for sequencing were prepared by adding 1 µl each of diluted enzyme, dNTP mixture and one ddNTP (G, A, T, or C) to a series of 4 tubes. Upon completion of the annealing reaction, 2 µl of the primer/template mixture was added to each of the 4 tubes and the extension-termination reaction allowed to proceed at 50°C for 30 min. Reactions were then stopped by adding 10 µl stopping mixture (98% formamide, 10 mM EDTA, 0.08% bromophenol blue and 0.08% xylene cyanol).

TABLE 2. Oligonucleotide primers.

Primer	Length (nt.)	Position[1]	Sequence[2]
# 1	18	403-420	TAGCCTTGGAGGATGGTC
# 2	17	1597-1613	GGGGTTTGGCTGTGTCC
# 3	16	1825-1840	CCTTCCGGCACCGGGC
# 4	17	2403-2419	CCCGGTAGCGAGTTGCC
# 5	17	2875-2891	GGTCAAGCCTCACGGGC
# 6	17	968-984	GTCTGGGTTGTTTCCCT
# 7	18	2383-2400	CCATACAGAACCACCGGA
# 8	17	1580-1596	TACGGTTCCTGGTTACC

1 Hybridizing locations of the oligonucleotides in the *P. aeruginosa* 23S rRNA in the 5'–>3' orientation; the primers anneal to these sites so that their 5'-ends anneal to the 3'-end of the RNA sequence and nucleotides are incorporated at the 3'-end of the primer.

2 Sequences listed are those of the synthesized primers in the 5'—>3' direction.

Separation of fragments.

Aliquots of 1-2 µl of the quenched reaction mixtures were separated on 6% polyacrylamide gels (+ 8.3 M urea) with 0.6 x TEB as running buffer. Just before they were loaded on the gel, the reaction mixtures were heated at 90°C for 1 min and plunged in ice water to ensure their denaturation. Two gels (length 85 cm and thickness 0.2 mm) were run to obtain sequences of lengths up to 250-300 nt from each reaction: the first run was stopped when the bromophenol blue reached the bottom (after roughly 3 hours at 75 watts), and the second run was allowed to continue for 7-8 hours, i.e. for 2-3 hours after the xylene cyanol had run off the bottom. The gels were wrapped in Saran Wrap and autoradiographed at -20°C.

Data manipulation: sequence alignment and construction of phylogenetic trees.

All alignments were done using the 23S rRNA sequence of *P. aeruginosa* (Gutell and Fox, 1988) as the index for comparison. Sequences were aligned by eye, using the Monaco font in the Microsoft Word program to ensure uniformity in all the characters (the different alphabets and numerals vary proportionally in the conventional Times font, while in the Monaco font all characters occupy the same width. e.g. G g A a). Sequences were first aligned with respect to conserved primary structure, and in highly variable regions secondary structural features (stem-loop structures) were used as markers. Only unambiguously alignable portions of the molecule were considered for comparative analyses (Lane *et al.*, 1985).

Comparisons were conducted on a pairwise basis, and the dissimilarity between two sequences determined by counting mismatched residues. Single position gaps were counted as being equal to a single substitution (Hori and Osawa, 1979). The differences were expressed as fractions of the mismatches between two

organisms. K_{nuc} values were calculated from these mismatch fractions and used to construct dendrograms.

RESULTS

Restriction fragment length analysis

The banding patterns obtained by digesting genomic DNAs of *Pseudomonas* species with the restriction enzymes *Eco*RI, *Eco*RV and *Hin*dIII are displayed in Figs. 1, 2 and 3 respectively. The bulk of the fragments range from 2-20 kb as indicated by the markers (λ DNA digested with *Cla*I and *Bst*EII). While there are discernible differences in the banding patterns of the various species, the recognition of specific, individual patterns is not possible because of the high numbers of fragments generated.

rRNA hybridizations

Autoradiograms of Southern transfers derived from the above analyses and hybridized against *Pseudomonas* rp2 rRNA are shown in Figs. 4, 5, and 6. The differences in patterns among the various species reflect variations in the flanking sequences of the rRNA operons. In each lane 5 or 6 individual bands of roughly equal intensities are clearly visible, implying that there are at least as many rRNA gene clusters present in each genome. The presence of strong low molecular weight bands indicates the presence of restriction sites within the rRNA operon. The *Lysobacter* rRNA gene cluster has sites susceptible to all three enzymes and *Hin*dIII sites are present in the rRNA operons of all the *Pseudomonas* species. Distinct hybridization patterns are clearly identifiable among the different strains. *Pseudomonas* rp2 and *P. viridiflava* show identical patterns when digested with *Eco*RI. Common bands are observable in these two lanes in the other digests also. Several common bands can be seen amongst the organisms *Pseudomonas aureofaciens*, *Pseudomonas fluorescens*, *Pseudomonas marginalis* and the field isolates *Pseudomonas* sp (wheat) and *Pseudomonas* sp (canker).

Figure 1. Agarose gel electrophoretic separation of *Eco*RI digests of genomic DNAs. DNA fragments were visualized by staining with ethidium bromide and illuminating with UV light.

Lanes: B: *Bst*EII digested λDNA; 1: *L. enzymogenes*; 2: *P. aureofaciens*; 3:*Pseudomonas*. sp (canker); 4: *P. cichorii*; 5: *P. fluorescens*; 6: *P. marginalis*; 7: *P. syringae* pv. *tabaci*; 8: *P. syringae* pv..*maculicola*; 9: *P. syringae* pv. *syringae*; 10: *Pseudomonas* rp2; 11: *P. viridiflava*; 12: *Pseudomonas*. sp (wheat); C: *Cla*I digested λDNA.

Markers: lane B: The sizes of the markers indicated by pointers (from top to bottom) are: 14,140 bp, 8454 bp, 6369 bp, 3675 bp, 2323 bp and 1264 bp. The remaining markers are 7242 bp, 5686 bp, 4822 bp, 4324 bp, 1929 bp and 1371 bp. lane C: The sizes of selected markers are indicated. The remaining markers are: 10,469 bp, 4398 bp, 2614 bp, and 1804 bp.

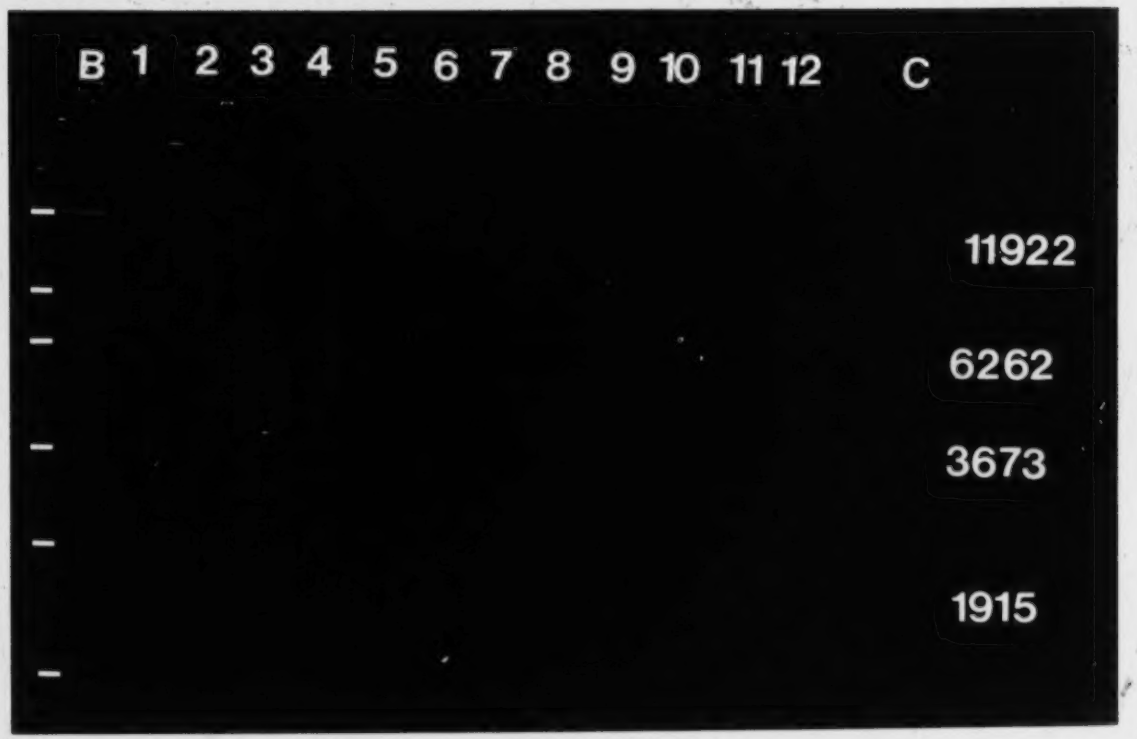

B 1 2 3 4 5 6 7 8 9 10 11 12 C
11922
6262
3673
1915

Figure 2. Agarose gel electrophoretic separation of *Eco*RV digests of genomic DNAs. Lanes and markers are as in Figure 1 with the following exceptions: 7: *Pseudomonas* rp2 and 10: *P. syringae* pv. *tabacii*..

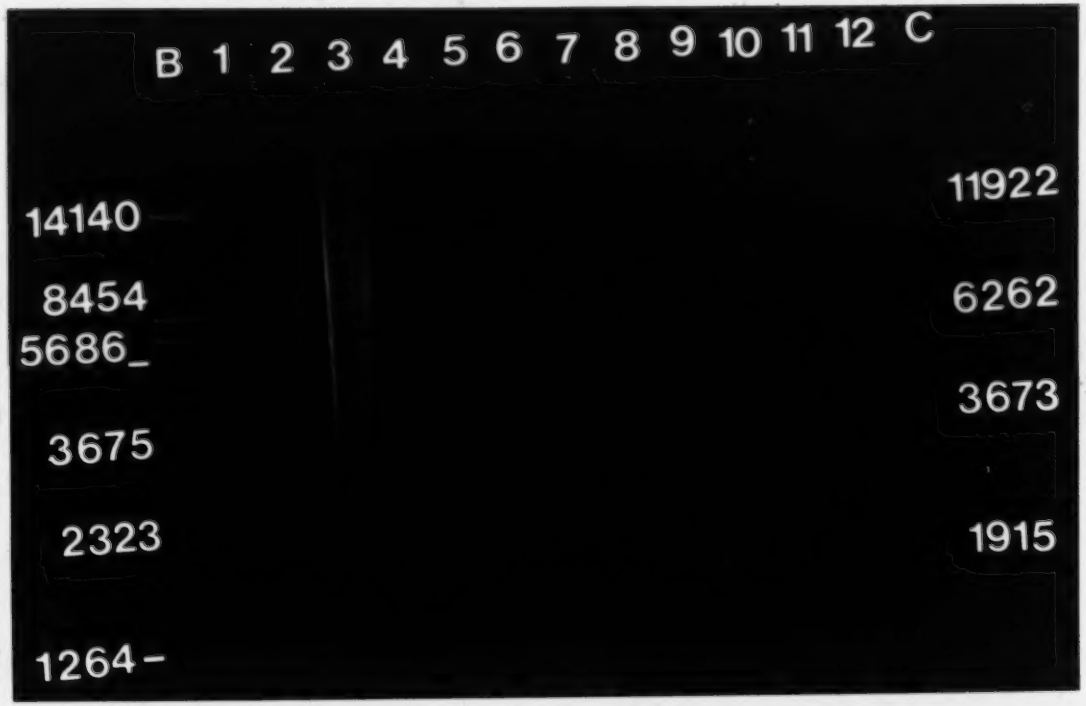

B 1 2 3 4 5 6 7 8 9 10 11 12 C
14140
8454
5686_
3675
2323
1264-
11922
6262
3673
1915

Figure 3. Agarose gel electrophoretic separation of *Hin*dIII digests of genomic DNAs. Markers and lanes are as in Figure 1 with the following exceptions: 7: *Pseudomonas* rp2 and 10: *P. syringae* pv. *tabacii*..

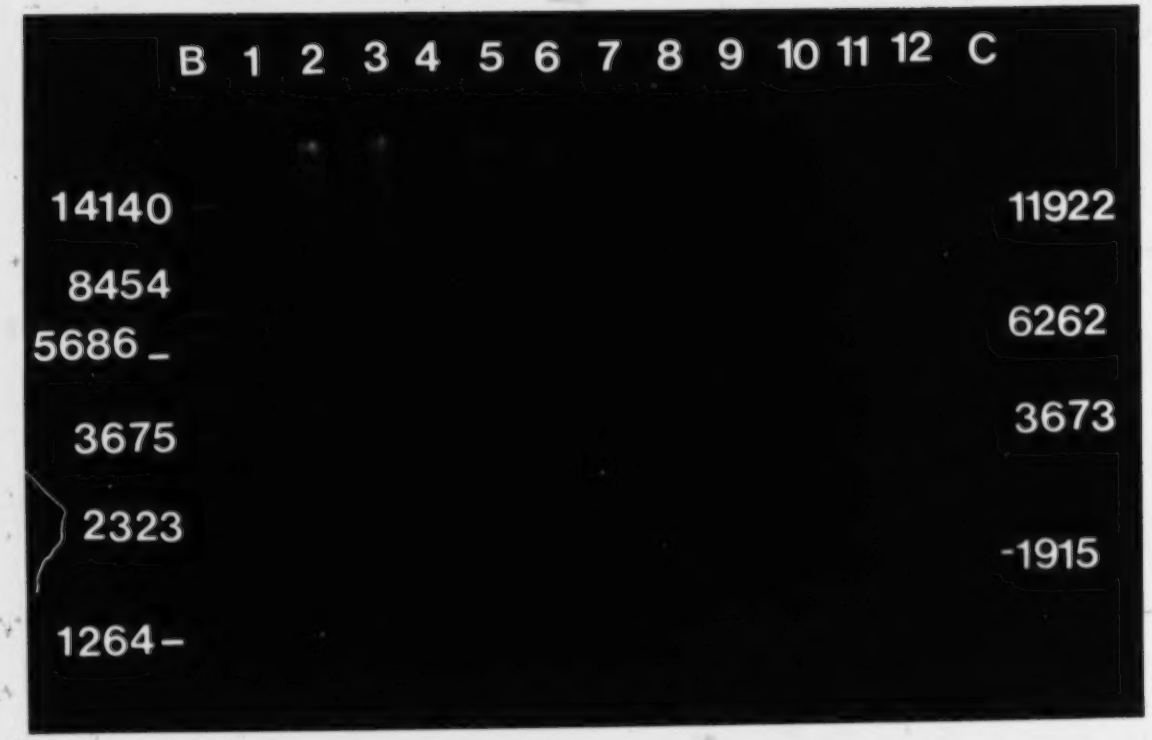

B 1 2 3 4 5 6 7 8 9 10 11 12 C
14140
11922
8454
6262
5686
3675
3673
2323
-1915
1264

Figure 4. Autoradiograph of the genomic DNAs digested with *Eco*RI and probed with *Pseudomonas* rp2 rRNA. Lanes: 1: *L. enzymogenes*; 2: *P. aureofaciens*; 3:*Pseudomonas*. sp (canker); 4: *P. cichorii*; 5: *P. fluorescens*; 6: *P. marginalis*; 7: *P. syringae* pv. *tabaci*; 8: *P. syringae* pv. *maculicola*; 9: *P. syringae* pv. *syringae*; 10: *Pseudomonas* rp2; 11: *P. viridiflava*; 12: *Pseudomonas*. sp (wheat). The arrowhead indicates the weakly hybridizing band in lane 1.

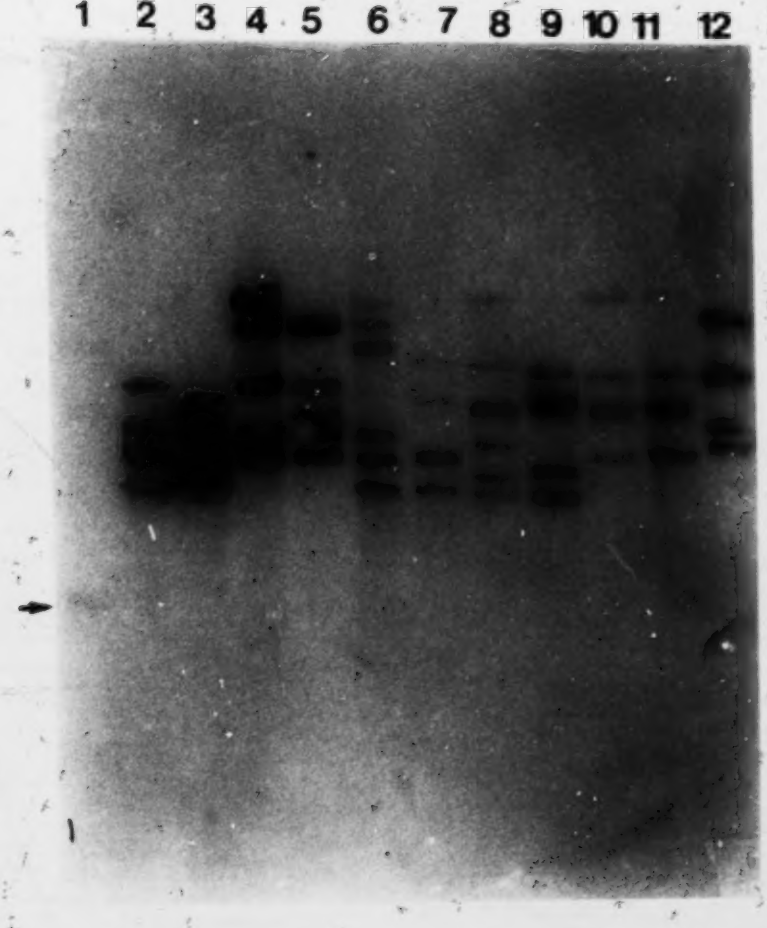

1 2 3 4 5 6 7 8 9 10 11 12

Figure 5. Autoradiogram of the genomic DNAs digested with *Eco*R V and probed with *Pseudomonas* rp2 rRNA. Markers and lanes are as in Figure 1 with the following exceptions: 7: *Pseudomonas* rp2 and 10: *P. syringae* pv. *tabacii*..

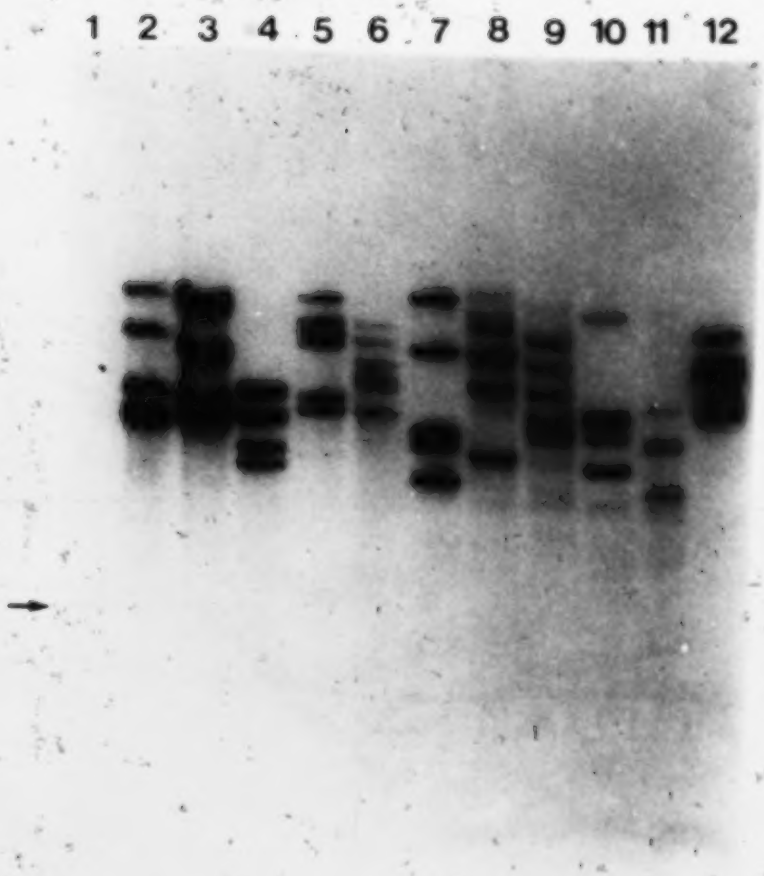

1 2 3 4 5 6 7 8 9 10 11 12

Figure 6. Autoradiographs of the genomic DNAs digested with *Hin*DIII and probed with *Pseudomonas* rp2 rRNA. Markers and lanes are as in Figure 1 with the following exceptions: 7: *Pseudomonas* rp2 and 10: *P. syringae* pv. *tabacii*..

1 2 3 4 5 6 7 8 9 10 11 12

23S rRNA sequencing

Fig. 7 depicts the 23S rRNA molecule from *P. aeruginosa* in its secondary structure format as proposed by Gutell and Fox (1988). The primer-hybridizing sites are indicated in the diagram. Reverse transcriptase-catalyzed sequencing was the method of choice for these experiments since the rRNA templates could be isolated directly from the cells in sufficient quantities for analyses. A frequent problem encountered with this technique was the occurrence of bands in all 4 lanes at certain positions on the gel (see Figure 8). Nucleotides at these positions were tentatively identified assuming correct termination in the lane with the maximum intensity. When all 4 lanes had similar intensities, the correct nucleotide at that position could not be determined. At such locations, sequence data from other isolates were compared to derive a consensus sequence for that region, based on the conservation of the sequence. The most likely nucleotide could then be predicted from this consensus.

The priming sites were chosen so that all six domains were represented. Of the first six oligonucleotides synthesized, two (#2 and #4) proved unsuccessful as primers for extension in all instances. Therefore oligonucleotides #7 and #8 were synthesized as alternatives. Their sequences extended into base paired regions just distal to the original primers, and they proved effective in initiating synthesis.

Figures 9-14 represent the alignment of partial sequences of the 23S rRNA from the authentic *Pseudomonas* spp. used in this study. Due to the high degree of similarity among all these organisms, sequences were quite easily aligned. Altogether, 950-1400 nucleotide positions were considered for comparative analyses. Positions of uncertain identity were not considered (Lane *et al.*, 1985). The rRNAs from all organisms demonstrate a high degree of homology to one another. The most distantly related organism appears to be *P. aeruginosa*, while all others fall within a very narrow range of differences.

Figure 7. Secondary structure of the complete 23S rRNA molecule of *P. aeruginosa*. (a): 5'-half of the molecule; (b): 3'-half of the molecule. The numbers in small type refer to the position of the nucleotides. The six domains are numbered I–VI according to the scheme of Leffers *et al.* (1987). Primer hybridizing sites are outlined and and identified by number. The tertiary interaction between two triplet sequences in domain V is also indicated. This figure has been reprinted from Figure 2, Gutell and Fox, (1988); with the permission of the editors of *Nucleic Acids Research*, IRL Press.

Figure 7a & b: Secondary structure of the complete 23S rRNA molecule of *P. aeruginosa*

reproduced from: R.R. Gutell and G.E. Fox (1988). Compilation of large ribosomal subunit RNA sequences presented in a structural format. Nucleic Acids Research 16 (sequence supplement) 175-270.

has been removed pending receipt of copyright permission.

Figure 8. Autoradiograph of a representative reverse transcriptase sequencing gel. A portion of a typical 6% polyacrylamide/8.3 M urea sequencing gel is shown. The specific example is that of *P. s. tabaci*, using primer #5. Lanes containing reactions terminated with the specific ddNTP are identified at the top and the actual sequence is shown on the right. This sequence is complementary to the rRNA sequence reported in the alignments. Distances from the 3'-end of the primer are indicated at the upper and lower left corners of the gel. Non-specific terminations are specified by bold arrows on the left.

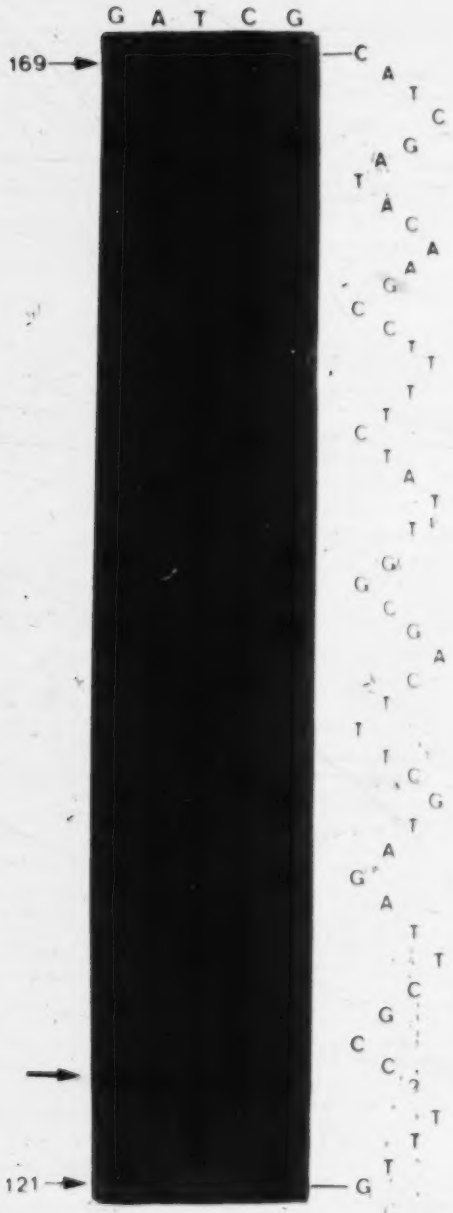

G A T C G
169
121

Figure 9. Partial sequence alignment of the authentic *Pseudomonas* species with primer #1. All alignments were done with the sequence of *P. aeruginosa* as the index for comparison. The sequence for this organism is displayed, along with the position of the first nucleotide in every line. The sequence determined by primer #1 lies in domain I and stretches from approximately 150-402.

- represents positions identical to the reference strain:

Upper case letters denote actual changes at a given position;

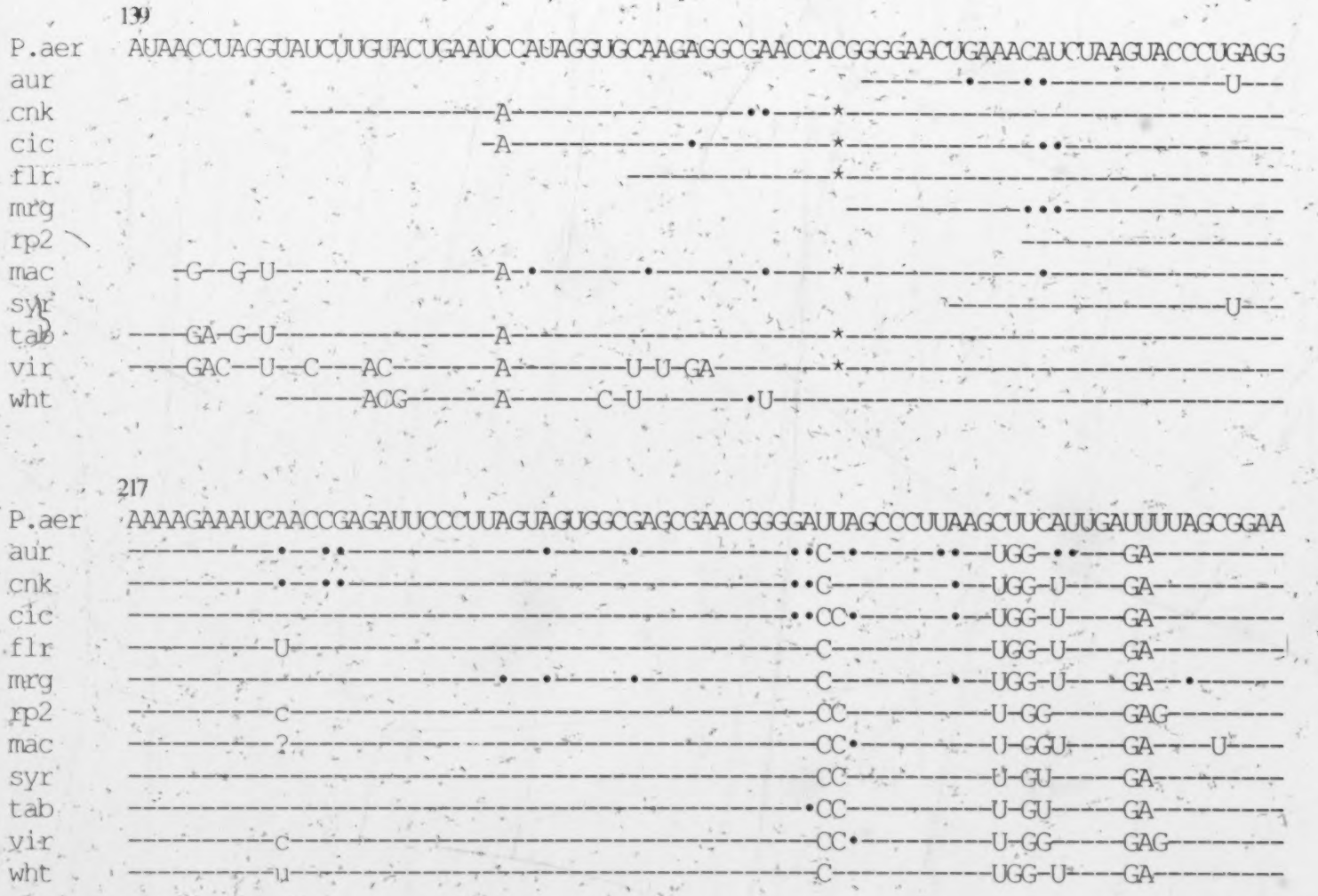
* indicates a gap.

The symbol • and lower case letters indicate positions which could not be determined with absolute certainty (see text for details).

The ? is used to indicate positions where no predictions could be made regarding the nucleotide most likely to occupy that site.

_____ : primer annealing site.

The abbreviations used are: P.aer: *P. aeruginosa*; aur: *P. aureofaciens*; cnk: *P. sp* (canker); cic: P. cichorii; flr: P. *fluorescens*; mrg: *P. marginalis*; rp2: *Pseudomonas* rp2; mac: *P. syringae* pv.*maculicola*; syr: *P. syringae* pv. *syringae*; tab: *P. syringae* pv. *tabaci*; vir: *P. viridiflava*; wht: *P. sp.* (wheat).


139
P.aer AUAACCUAGGUAUCUUGUACUGAAUCCAUAGGUGCAAGAGGCGAACCACGGGGAACUGAAACAUCUAAGUACCCUGAGG
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht
217
P.aer AAAAGAAAUCAACCGAGAUUCCCUUAGUAGUGGCGAGCGAACGGGGAUUAGCCCUUAAGCUUCAUUGAUUUUAGCGGAA
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht

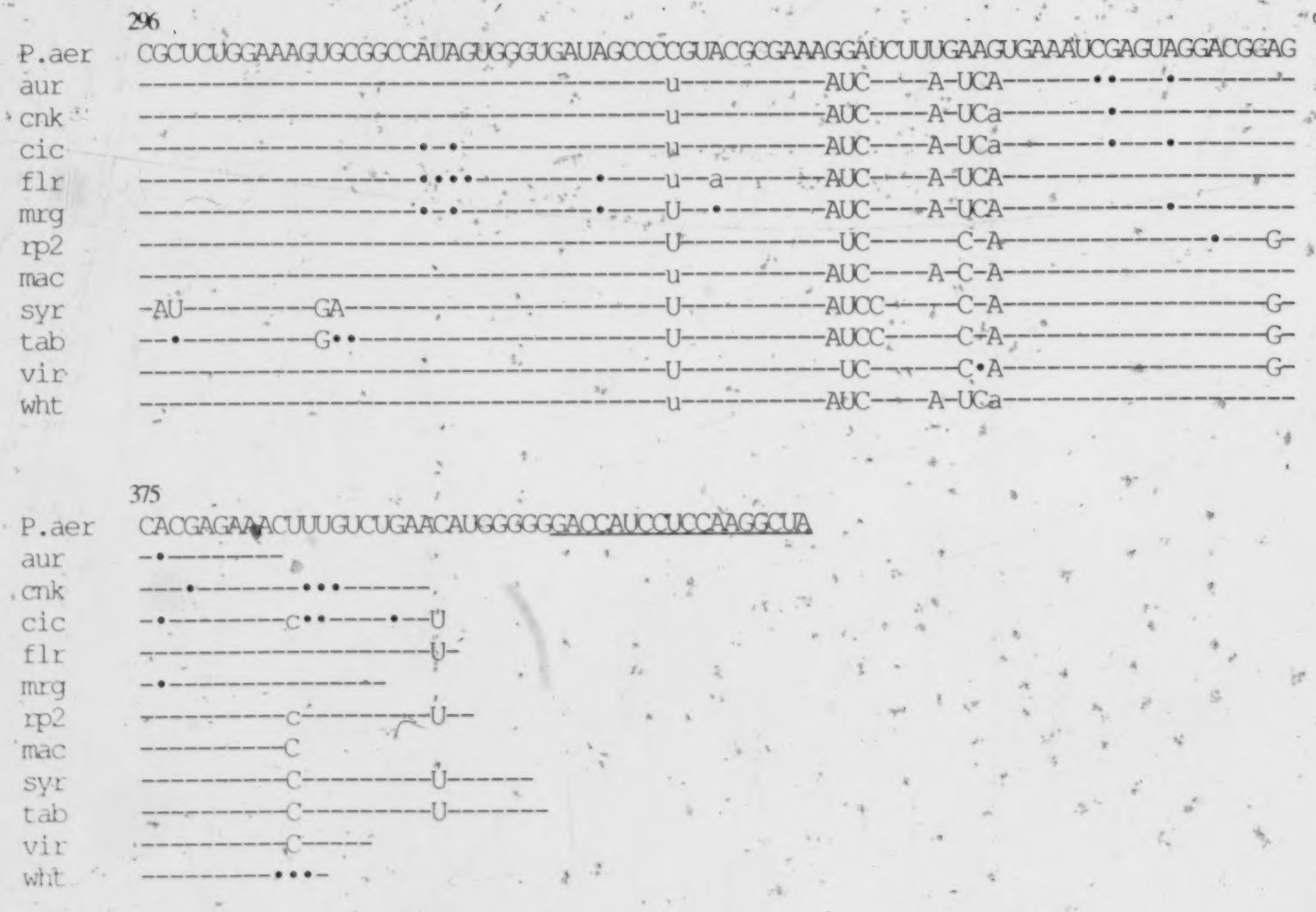


Figure 9 (continued).

Figure 10. Partial sequence alignment of the authentic *Pseudomonas* species with primer #6. This alignment includes the region of the molecule from 740-967 in domain II. All the symbols and abbreviations used are as described in the legend to Fig. 9.

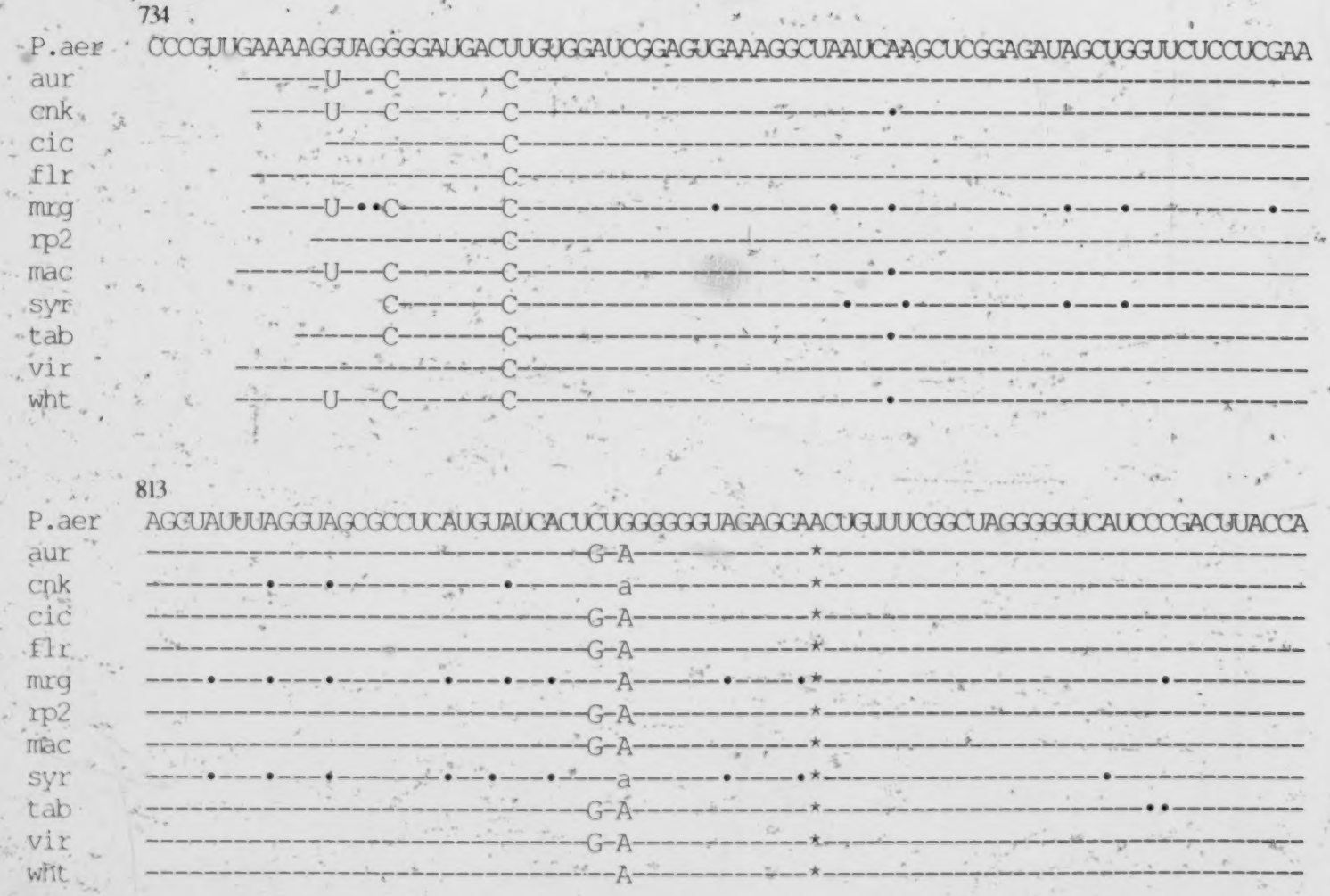

734
P.aer CCCGUUGAAAAGGUAGGGGAUGACUUGUGGAUCGGAGUGAAAGGCUAAUCAAGCUCGGAGAUAGCUGGUUCUCCUCGAA
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht
813
P.aer AGCUAUUUAGGUAGCGCCUCAUGUAUCACUCUGGGGGGUAGAGCAACUGUUUCGGCUAGGGGGUCAUCCCGACUUACCA
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht

Figure 10 (continued).

```
                                                                  892
P.aer   AACCGAUGCAAACUCCGAAUACCCAGAAGUGCCGAGCAUGGGAGACACGGCGGGUGCUAACGUCCGUCGUGAAAGG

aur     ----------------------U-C-----------------------------------------------------
cnk     ----------•-----------U-----•-----------------------------------------------
cic     ----------------------U-C-----------------------------------------------------
flr     ----------------------U-C-----------------------------------------------------
mrg     ----------•-----------U-----•-----•-----•-----•-----•-----•-----•-----•-----•
rp2     ----------------------U-C-----------------------------------------------------
mac     ----------------------U-C-----•-----------------------------------------------
syr     ----------•-----•-----U-----•-----•-----•-----•-----•-----•-----•-----•-----•
tab     ----------------------U-C-----•-----•-----------------------------------------
vir     ----------------------U-C-----------------------------------------------------
wht     ----------------------U-C-----•-----•-----•-----•-----•-----•-----•-----•-----

P.aer   CAAACAACCCAGAC
```

Figure 11. Partial sequence alignment of the authentic *Pseudomonas* species with primer #8. Sequences in this figure range from 1329-1579. All symbols and abbreviations are as described in the legend to Fig. 9.

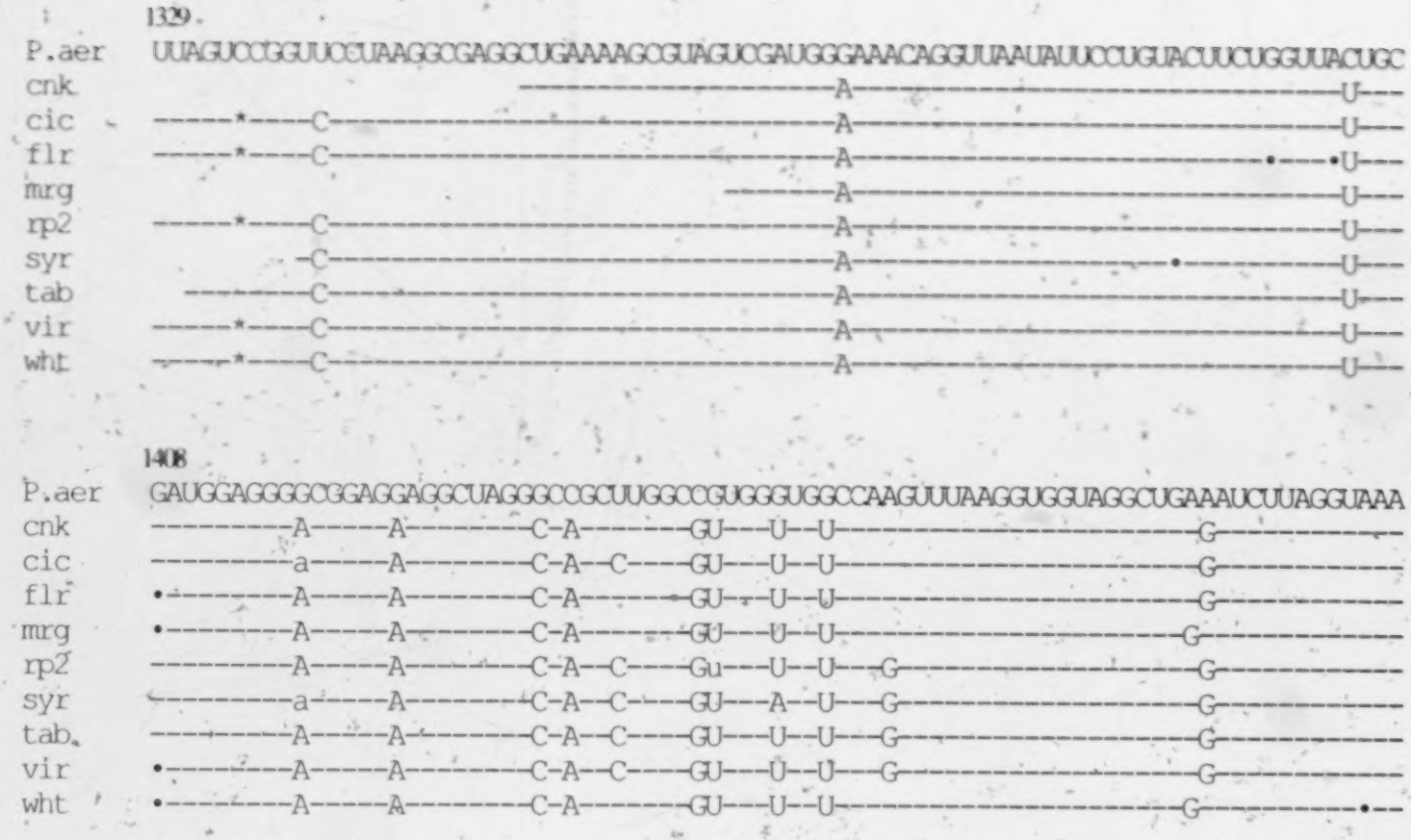

1329
P.aer UUAGUCCGGUUCCUAAGGCGAGGCUGAAAAGCGUAGUCGAUGGGAAACAGGUUAAUAUUCCUGUACUUCUGGUUACUGC
cnk
cic
flr
mrg
rp2
syr
tab
vir
wht
1408
P.aer GAUGGAGGGGCGGAGGAGGCUAGGGCCGCUUGGCCGUGGGUGGCCAAGUUUAAGGUGGUAGGCUGAAAUCUUAGGUAAA
cnk
cic
flr
mrg
rp2
syr
tab
vir
wht

Figure 11 (continued).

```
                                                                1487
P.aer  UCCGGGGUUUCAAGGCCGAGAGGUGAUGACGAGUCGUCUUUUAGAUGACGAAGUGCUUGAUGCCAUGCUUCCAAGAAAA
cnk    ------A-C-U----------C----------UAA----------U-A----------G----------•--
cic    ------A-C-U----------C----------GU---**-----AC---------G----------•--
mrg    ------A--CU----------C----------UAA----------a-U-A--------uG---------•--
flr    ------A-C-U----•-----C----------GU----------AC---•------•G----------•--
rp2    ------A-C•-•---------C----------GU----------AC---------G----------•--
syr    ------A-C-U•----•----C----------Gu•---------AC-U-------G----------•--
tab    ------A-C----------C----------GU---*-------AC---------G----------•--
vir    ------A-C-U----------C----------GU---------G--AC-------G----------?--
wht    ------a--CU----------C----------U•---------G----------G----------•--

                                                                1566
P.aer  GCUUCUAAGCUUCAGGUAACCAGGAACCGUA
cnk    ----------
cic    ----------
flr    ----------
mrg    ----------
rp2    ----------
syr    ----------
tab    ----------
vir    ----------
wht    ----------
```

Figure 12. Partial sequence alignment of the authentic *Pseudomonas* species with primer #3. Sequences in this figure range from 1585-1824. All symbols and abbreviations are as described in the legend to Fig. 9. _____ indicates the hybridization site of primer #8, which overlaps into this region.

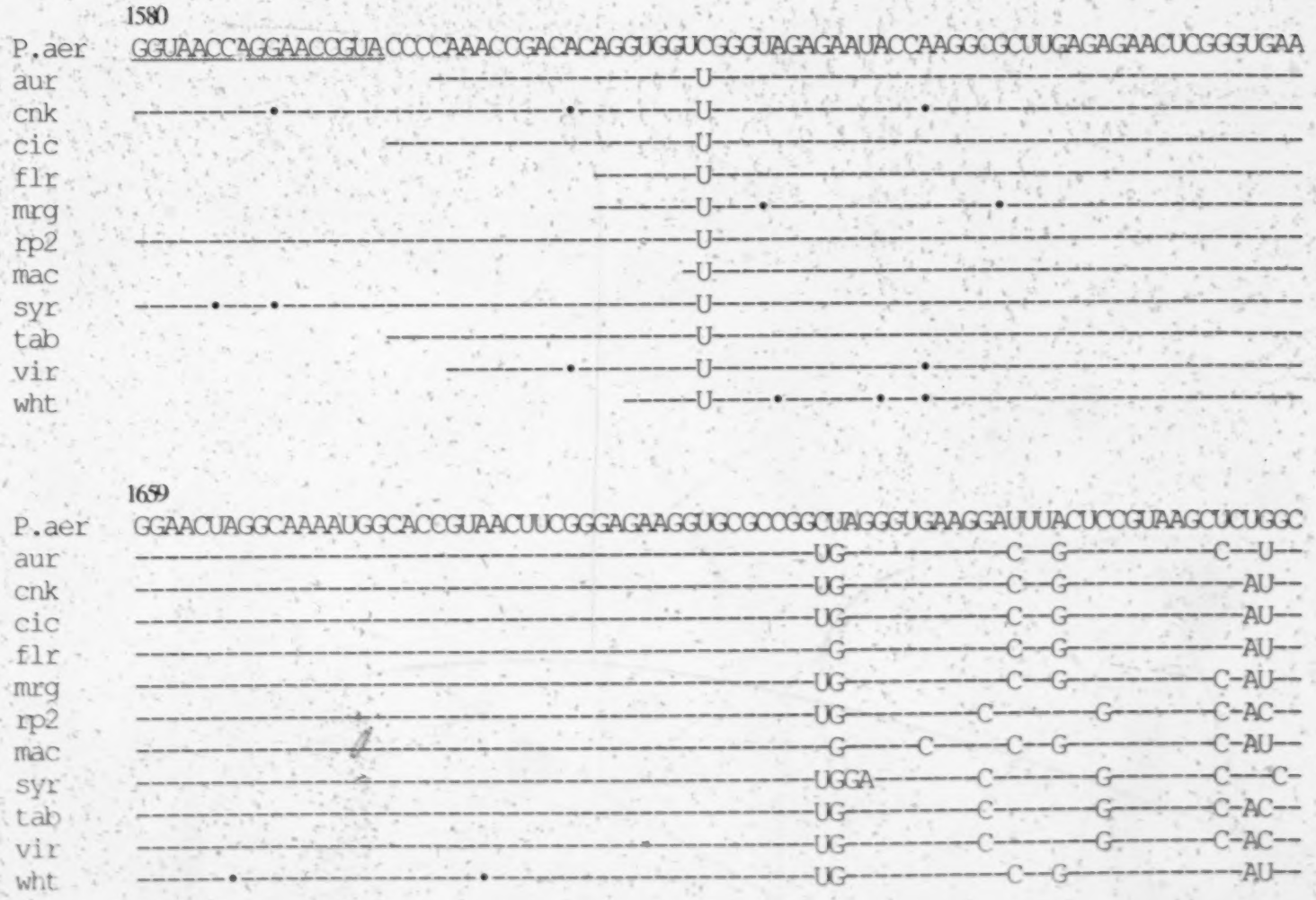

1580
P.aer GGUAACCAGGAACCGUACCCCAAACCGACACAGGUGGUCGGCUAGAGAAUACCAAGGCGCUUGAGAGAACUCGGGUGAA
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht
1659
P.aer GGAACUAGGCAAAAUGGCACCGUAACUUCGGGAGAAGGUGCGCCGGCUAGGGUGAAGGAUUUACUCCGUAAGCUCUGGC
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht

Figure 12 (continued)

```
                1738
P.aer  UGGUCGAAGAUACCAGGCCGCUGCGACUGUUAUUAAAAACACAGCACUCUGCAAACACGAAAGUGGACGUAUAGGGUG
aur    C--U-----------------------------------------------------------------------
cnk    C-------------------------------------------------------------------------
cic    C-------------------------------------------------------------------------
flr    C-------------------------------------------------------------------------
mrg    C--•----•-----•••-----------------------------------------------------------
rp2    C-------------------------------------------------------------------------
mac    C-------------------------------------------------------------------------
syr    c•••-----•-----------------------------------------------------------------
tab    C••-----------•-------------------------------------------------------------
vir    C-------•-----•--•----------------------------------------------------------
wht    C---•-------•-----•-------------------------------•---•---------------------

                1817
P.aer  UGACGCCUGCCCCGGUGCCCGGAAGG
aur    ---
cnk
cic
flr
mrg
rp2
mac
syr
tab    --
vir
wht    --•--
```

Figure 13. Partial sequence alignment of the authentic *Pseudomonas* species with primer #7. sequences in this figure range from 2140-2382. All symbols and abbreviations are as described in the legend to Fig. 9.

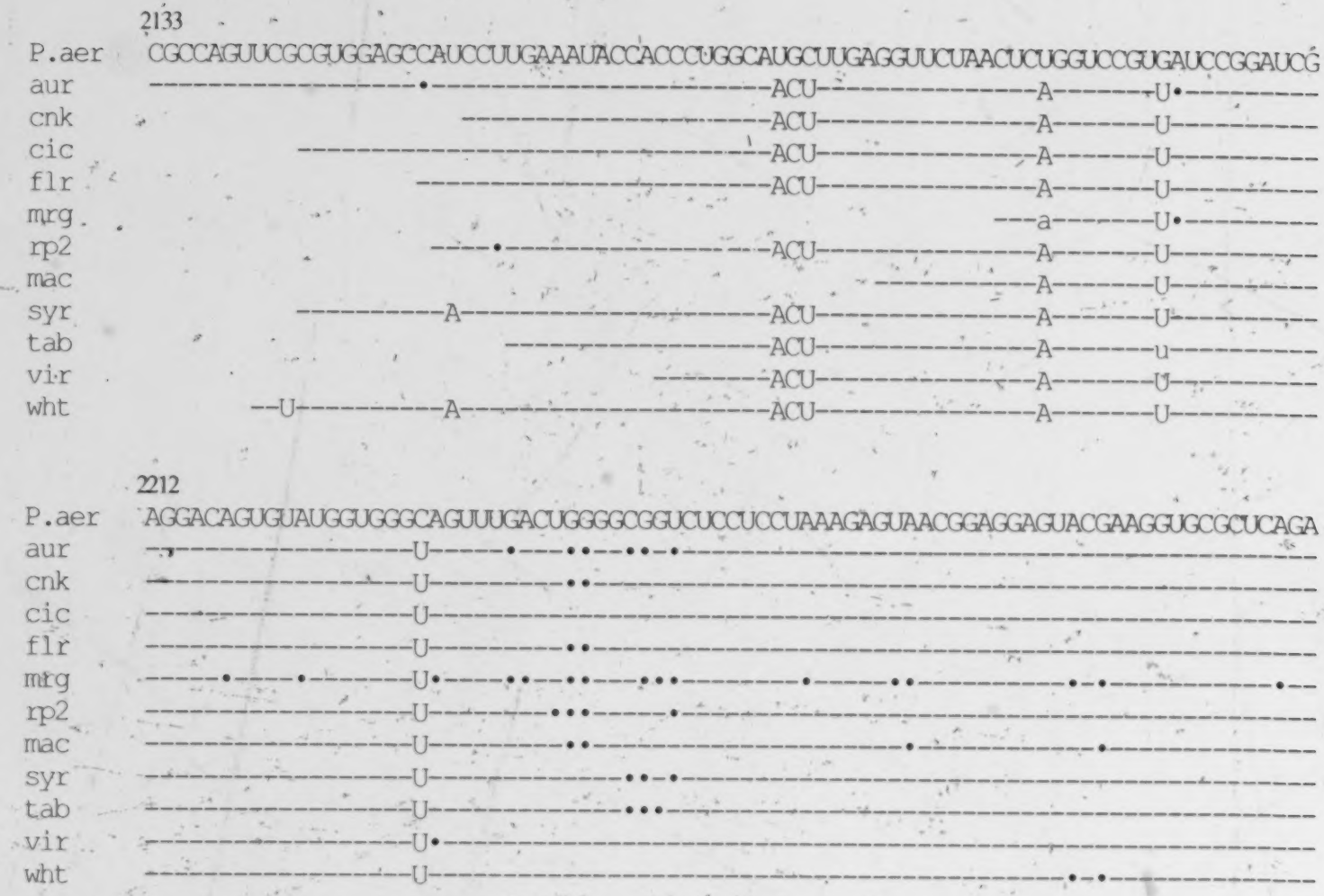

2133
P.aer CGCCAGUUCGCGUGGAGCCAUCCUUGAAAUACCACCCUGGCAUGCUUGAGGUUCUAACUCUGGUCCGUGAUCCGGAUCG
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht
2212
P.aer AGGACAGUGUAUGGUGGGCAGUUUGACUGGGGCGGUCUCCUCCUAAAGAGUAACGGAGGAGUACGAAGGUGCGCUCAGA
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht

Figure 13 (continued).

```
                2291
P.aer           CCGGUCGGAAAUCGGUCGCAGAGUAUAAAGGCAAAAGCGCGGUUGACUCCGAGACAGACACGUCGAGCAGGUACGAAAG
aur             ----------U----•-•----------•-----------------------•-----------•-•----------
cnk             ----------U----•-•----•-----•-----------------------•-----------•-•----------
cic             ----------U----•-•----•-----•-----------------------•-----------•-•----------
flr             ----------U----•-•----•-----•-----------------------•-----------•-•----------
mrg             ----•••••-U----•-•••••-----•-----------------------•-----------•-•----------
rp2             ----------U----•-•----•-----•-----------------------•-----------•-•----------
mac             ----------u----•-•----•-----•-----------------------•-----------•-•----------
syr             ----------U----•-•----•-----•-----------------------•-----------•-•----------
tab             ----------u----•-•----•-----•-----------------------•-----------•-•----------
vir             ----------U----•-•----•-----•-----------------------•-----------•-•----------
wht             ----------U----•-•----•-----•-----------------------•-----------•-•----------

                2370
P.aer           UAGGUCUUAGUGAUCCGGUGGUUCUGUAUGG
aur             --•-•----
cnk             --•-•----
cic             --•--
flr             --------
mrg             --•--
rp2             --------
mac             --•--
syr             --------
tab             --------
vir             --------
wht             --•---•--
```

Figure 14. Partial sequence alignment of the authentic *Pseudomonas* species with primer #5. Sequences in this figure range from 2616-2874. All symbols and abbreviations are as described in the legend to Fig. 9.

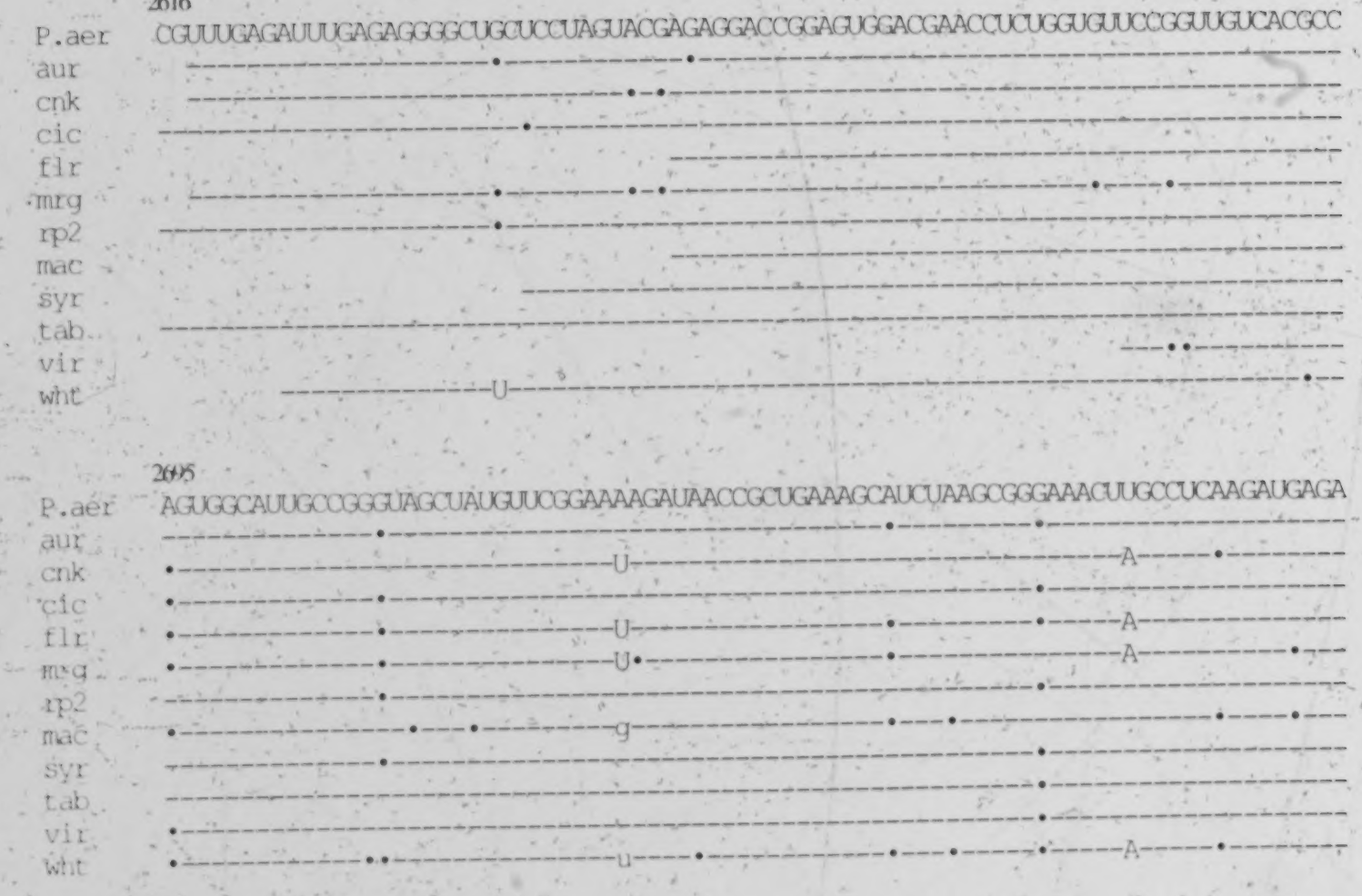

2616
P.aer CGUUUGAGAUUUGAGAGGGGCUGCUCCUAGUACGAGAGGACCGGAGUGGACGAACCUCUGGUGUUCCGGUUGUCACGCC
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht
2695
P.aer AGUGGCAUUGCCGGGUAGCUAUGUUCGGAAAAGAUAACCGCUGAAAGCAUCUAAGCGGGAAACUUGCCUCAAGAUGAGA
aur
cnk
cic
flr
mrg
rp2
mac
syr
tab
vir
wht

Figure 14 (continued).

The percent dissimilarities and K_{nuc} values for these organisms are summarized in Table 3.

Figs. 15-19 show alignments among more distant eubacterial species. The sequences of six of these were obtained from the literature (Gutell and Fox, 1988; Höpfl, 1989), while partial sequences of *L. enzymogenes* and *P. acidovorans* were determined as detailed before. *P. acidovorans* showed maximum homology to *P. cepacia* (~89%). All of the gram negative organisms were closer to each other than to the gram positive bacteria (see Table 4). *L. enzymogenes* showed maximum homology to *P. aeruginosa*.

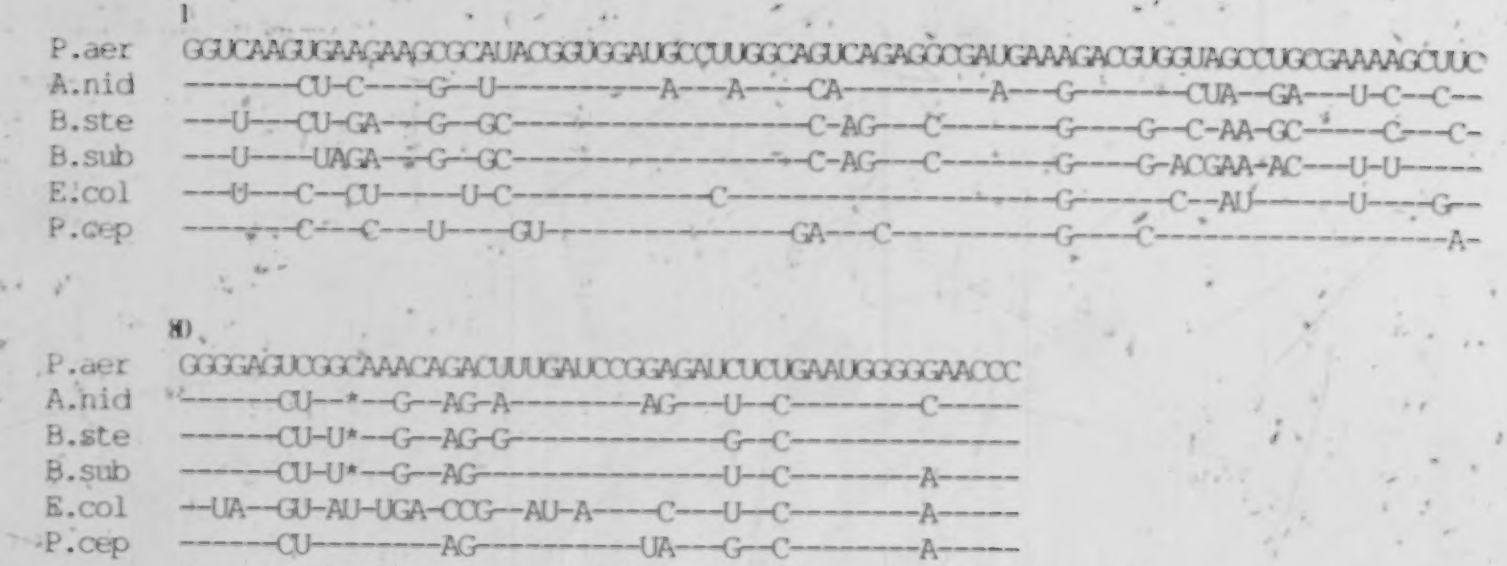
The evolutionary trees shown in Fig 20 and 21 were inferred from the K_{nuc} values which were estimated for each pair of organisms by the formula $-0.75 \times \mathrm{Ln}(1-4/3\,\lambda)$, where λ is the proportion of differences between the 2 organisms (Hori, 1975). Unrooted trees were plotted to avoid erroneous assumptions regarding the ancestral sequence. Fig 20 demonstrates the high degree of homology among the authentic *Pseudomonas* species. All the organisms tested form a cluster close to but distinct from *P. aeruginosa*. *Pseudomonas* rp2 is virtually identical to *P. viridiflava*, while the other 2 field isolates cluster with *P. aureofaciens*, *P. fluorescens*, and P. *marginalis*. The two pathovars *P. syringae* pv. *syringae* and *P. syringae* pv. *tabaci* appear to be somewhat closer to each other than to the other organisms but *P. syringae maculicola* does not show any such clustering.

TABLE 3. Pairwise comparison of the authentic *Pseudomonas* species.

Percentage dissimilarities (lower left) and K_{nuc} values (upper right) for each of the pairs of the authentic *Pseudomonas* spp studied are displayed below. These values were calculated from the sequence data shown in figs. 9-14. The organism names are abbreviated as described in the legend to fig. 9.

	P.aer	aur	cnk	cic	flr	mrg	rp2	mac	syr	tab	vir	wht
P.aer		0.043	0.051	0.057	0.054	0.053	0.052	0.048	0.06	0.058	0.065	0.052
aur	4.23		0.009	0.010	0.012	0.010	0.018	0.019	0.027	0.016	0.020	0.008
cnk	4.89	0.89		0.016	0.011	0.008	0.024	0.019	0.024	0.024	0.031	0.013
cic	5.50	0.96	1.55		0.012	0.024	0.017	0.020	0.029	0.019	0.021	0.018
flr	5.25	1.20	1.1	1.24		0.015	0.021	0.016	0.032	0.023	0.025	0.012
mrg	5.15	1.04	0.76	2.39	1.46		0.033	0.017	0.037	0.033	0.035	0.010
rp2	5.02	1.82	2.35	1.67	2.03	3.24		0.022	0.019	0.008	-	0.025
mac	4.69	1.91	1.83	1.99	1.57	1.69	2.13		0.031	0.021	0.031	0.023
syr	5.73	2.63	2.39	2.82	3.14	3.65	1.87	2.99		0.012	0.018	0.033
tab	5.59	1.55	2.32	1.88	2.23	3.24	0.76	2.03	1.16		0.016	0.033
vir	6.21	1.96	3.04	2.12	2.45	3.45	0.08	3.07	1.79	1.61		0.034
wht	5.04	0.79	1.32	1.74	1.22	1.01	2.48	2.25	3.23	3.25	3.32	

Figure 15. Partial sequence alignments of selected eubacterial species with primer #1. The sequences used in this analysis were obtained from the literature except those of *L. enzymogenes* and *P. acidovorans*, whose parial sequences were determined experimentally. Symbols are as indicated in the legend to Fig. 9. Nucleotide positions 1-402 have been included in this alignment. Sequences for *L. enzymogenes* and *P. acidovorans* spanned a smaller region from ~155 to 402. The sequences from positions 131-148, 270- 292, and 336-362 have been displayed but not aligned due to the variable nature of these regions. The abbreviations used are: A.nid: *Anacystis nidulans*; B.ste: *B. stearothermophilus*; B.sub: *Bacillus subtilis*; E. col: *E. coli*:; L.enz: *L. enzymogenes*; P.acd: *P. acidovorans*; P.aer: *P. aeruginosa*; P. cep: *P. cepacia*.


1
P.aer GGUCAAGUGAAGAAGCGCAUACGGUGGAUGCCUUGGCAGUCAGAGGCGAUGAAAGACGUGGUAGCCUGCGAAAAGCUUC
A.nid -------CU-C----G--U---------A---A----CA---------A---G------CUA--GA---U-C--C--
B.ste ---U---CU-GA---G--GC-----------------C-AG---C-------G----G--C-AA-GC-----C---C-
B.sub ---U----UAGA---G--GC-----------------C-AG---C-------G----G-ACGAA-AC---U-U-----
E.col ---U---C--CU-----U-C-------------C-------------------G------C--AU------U----G--
P.cep ------C---C---U----GU-----------------GA---C----------G----C-----------------A-
80
P.aer GGGGAGUCGGCAAACAGACUUUGAUCCGGAGAUCUCUGAAUGGGGGAACCC
A.nid ------CU--*--G--AG-A--------AG---U--C--------C-----
B.ste ------CU-U*--G--AG-G------------G--C--------------
B.sub ------CU-U*--G--AG---------------U--C--------A-----
E.col --UA--GU-AU-UGA-CCG--AU-A----C---U--C--------A-----
P.cep ------CU--------AG----------UA---G--C--------A-----
131
P.aer ACCUAGGAUAACCUAGGU
A.nid CAUGU
B.ste ACUGUCCGUAAUGGGGCAGU
B.sub ACCACUCGUAAUGGAGUGGU
E.col AGUGUGUUUCGACACACU
P.cep ACUGCUUUUGGAGU

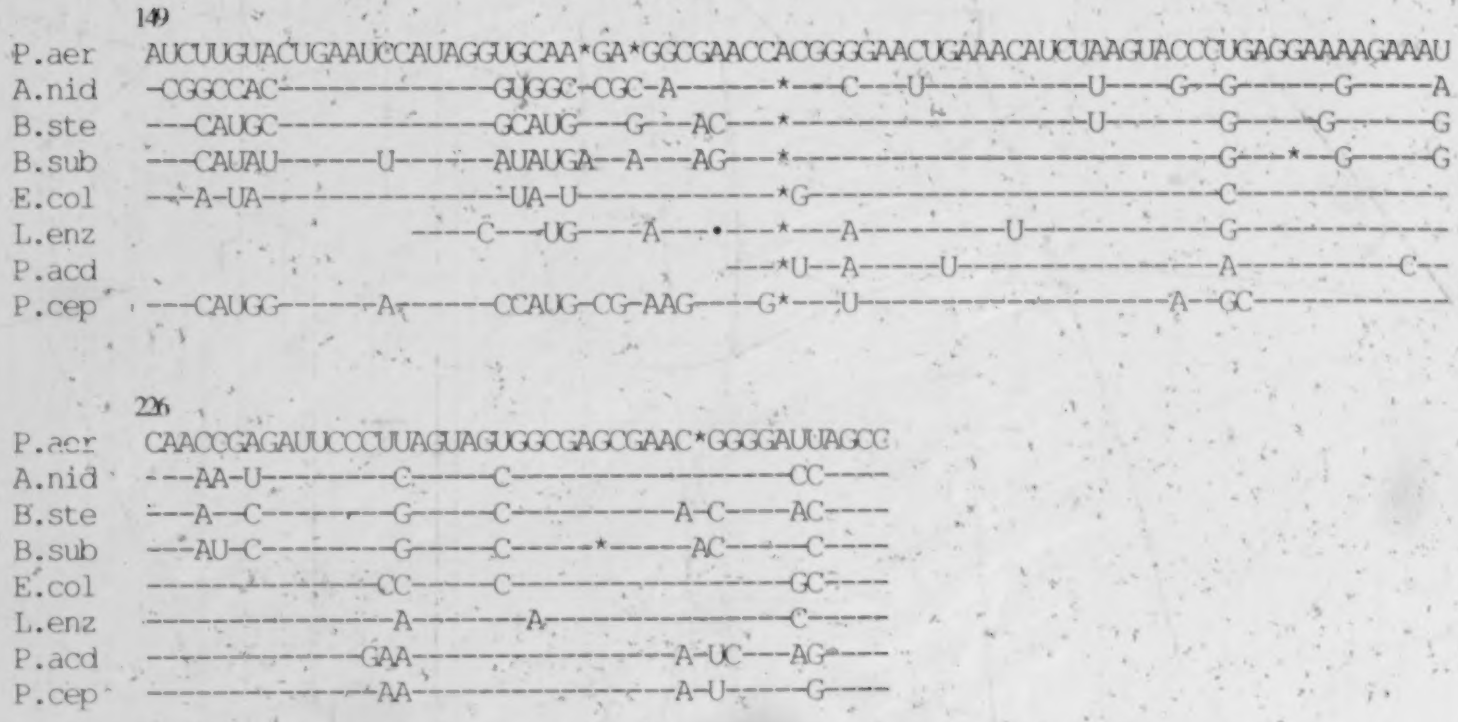


Figure 15 (continued).

```
          270
P.aer     CUUAAGCUUCAUUGAUUUUAGCG
A.nid     UAAACCAAACUCCACGGAGUUUGGGGUCGUGGGACAGCAAUGUGGACUGUGAAUGUUAGAC
B.ste     CAAACCAAGAGGCGUGCCUCUUGGGGUUGUAGGACCGCUCAUUGUGGGAGUGAGAAAGGAACGGGGUAGAC
B.sub     CAAACCAAGAGGCUUGCCUCUGUGGUUGUAGGACACUCUGUACGGAGUUACAAAAGAACGAGGUAGAU
E.col.    CAGAGCCUGAAUCAGUGUGUGUGUUAGUG
L.enz     CUUAAGUUGCAUCAGCUUCAGCA
P.acd     UUCUAGUGAUAGUCAGACGGUUAACA
P.cep     UUGCACUCUUUAUUUGUAUUGUUAGCC

          293
P.aer     GAAC*GCUCUGGAAAG*UGCGGCCAUAGUGGGUGAUAGCCCCCGUA
A.nid     ---G-CAG---*---A-CUGCA---G--AA-----A--U--U---
B.ste     ----C-G-------C--GC------G--AA-----C-----U---
B.sub     ---GA-G---------GGC-C-------GA---A-C-----U---
E.col     ---G-CG----------GCGC--G---CA------C---------
L.enz     A--G-AAC---------GUU----G---AA-----C-----U---
P.acd     A-----GAA---------C----------A----------U---
P.cep     ---------------------------CA----------U---
```

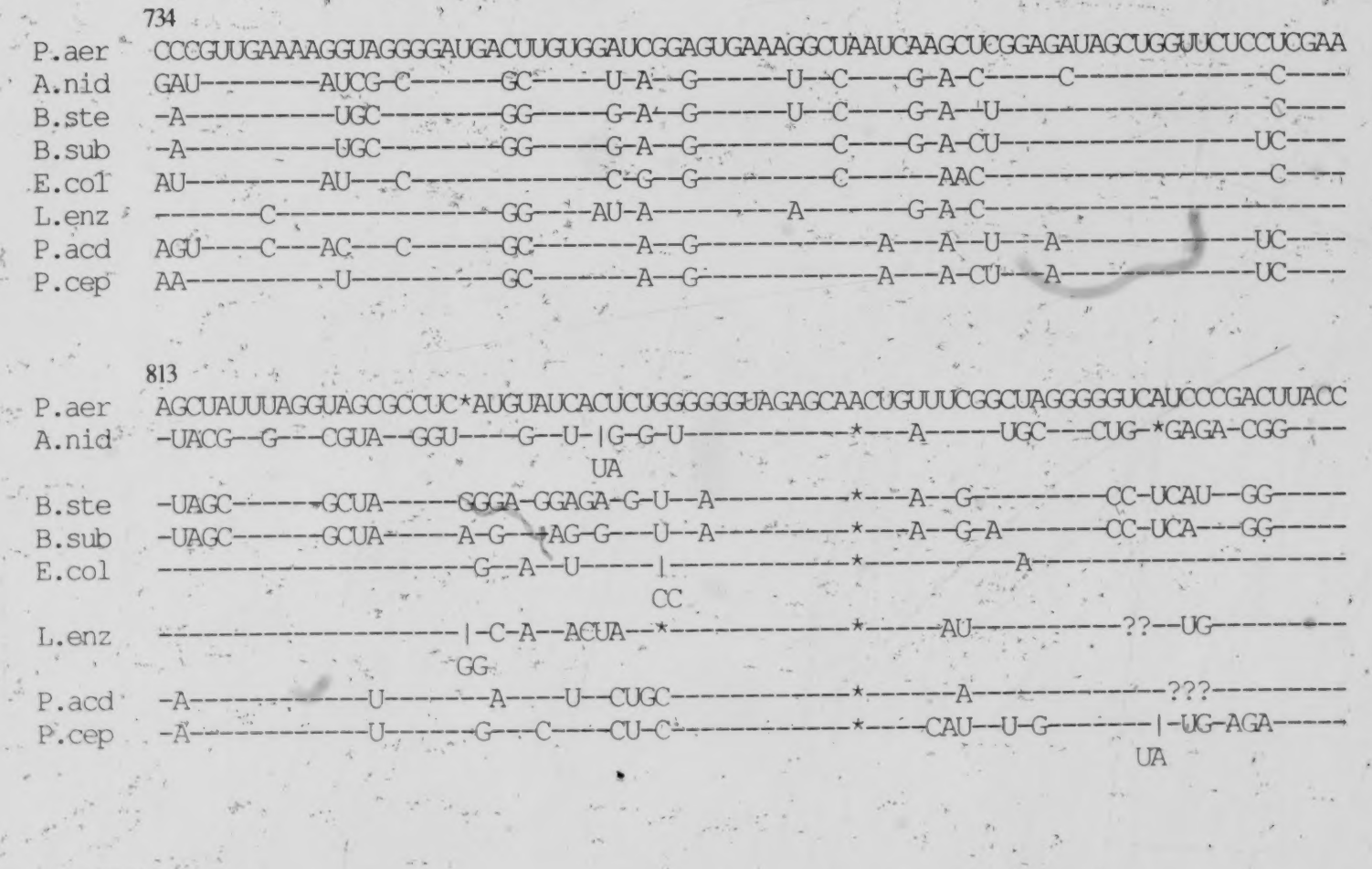
Figure 15 (continued).

Figure 15 (continued)

```
                336
P.aer  CGCGAAAGGAUCUUUGAAGUGAAAUCG
A.nid  GUCGAAAAUUGAAACAGCCUAGCUGAAUCCCG
B.ste  GUCGAAACUUCGUUCCCUCCCGAGCGGAUCCUG
B.sub  GUCAAAACUUCGUUCUCUCCUGAGUGGAUCCUG
E.col  CACAAAAAUGCACAUGCUGUGAGCUCGAUG
L.enz  UGCGAAGGGGCUGAUGC?AUGAAAUCG
P.acd  UGUGAAAACCGACUGGGUACUGAGCUAGAGAAA
P.cep  GGCGAAAACAGUAUGAAAGAACUAGGUGUGGCGACA

                364
P.aer  AGUAGGACGGAGCACGAGAAACUUUGUCUGAACAUGGGGGGACCAUCCUCCAAGGCUA
A.nid  ------C----------U----U-CC--G----UCC-C-A------C-UCGU------
B.ste  ------C-G----GA----G----U-CC--G----GCA----A-------UC------
B.sub  ------C-G-----A----U----U-CC--G----UCC----A-------UC------
E.col  ------G----GA----U-GU-UCC-----------U-----------------------
L.enz  ------G-•-G-----------CC?-----------------------------------
P.acd  ------G---G-----------CC?•----U-----------------------------
P.cep  ------G---GA----U----UCC-----G-----------------------------
```

Figure 16. Partial sequence alignments of selected eubacterial species with primer #6. Sequences range from positions 734-967. All symbols are as described in the legend to Fig. 9 and abbreviations as in Fig. 15..


734
P.aer CCCGUUGAAAAGGUAGGGGAUGACUUGUGGAUCGGAGUGAAAGGCUAAUCAAGCUCGGAGAUAGCUGGUUCUCCUCGAA
A.nid GAU--------AUCG-C------GC-----U-A--G------U--C----G-A-C-----C----------C----
B.ste -A----------UGC--------GG-----G-A--G------U--C----G-A--U----------------C----
B.sub -A----------UGC--------GG-----G-A--G---------C----G-A-CU----------------UC----
E.col AU----------AU---C-------------C-G--G---------C------AAC-----------------C----
L.enz -------C--------------GG---AU-A---------A-------G-A-C-----------------------
P.acd AGU----C---AC---C------GC-------A--G----------A---A--U---A-------------UC----
P.cep AA----------U----------GC-------A--G----------A---A-CU---A-------------UC----
813
P.aer AGCUAUUUAGGUAGCGCCUC*AUGUAUCACUCUGGGGGGUAGAGCAACUGUUUCGGCUAGGGGGUCAUCCCGACUUACC
A.nid -UACG--G---CGUA--GGU----G--U-|G-G-U----------*---A-----UGC---CUG-*GAGA-CGG----
UA
B.ste -UAGC------GCUA-----GGGA-GGAGA-G-U--A---------*---A--G---------CC-UCAU--GG-----
B.sub -UAGC------GCUA-----A-G---AG-G---U--A---------*---A--G-A-------CC-UCA---GG-----
E.col ----------------------G--A--U-----|----------*----------A--------------------
CC
L.enz ---------------------|-C-A--ACUA--*----------*-----AU----------??--UG---------
GG
P.acd -A----------U-------A----U--CUGC----------*------A----------???---------
P.cep -A----------U------G---C----CU-C----------*----CAU--U-G-------|-UG-AGA-----
UA

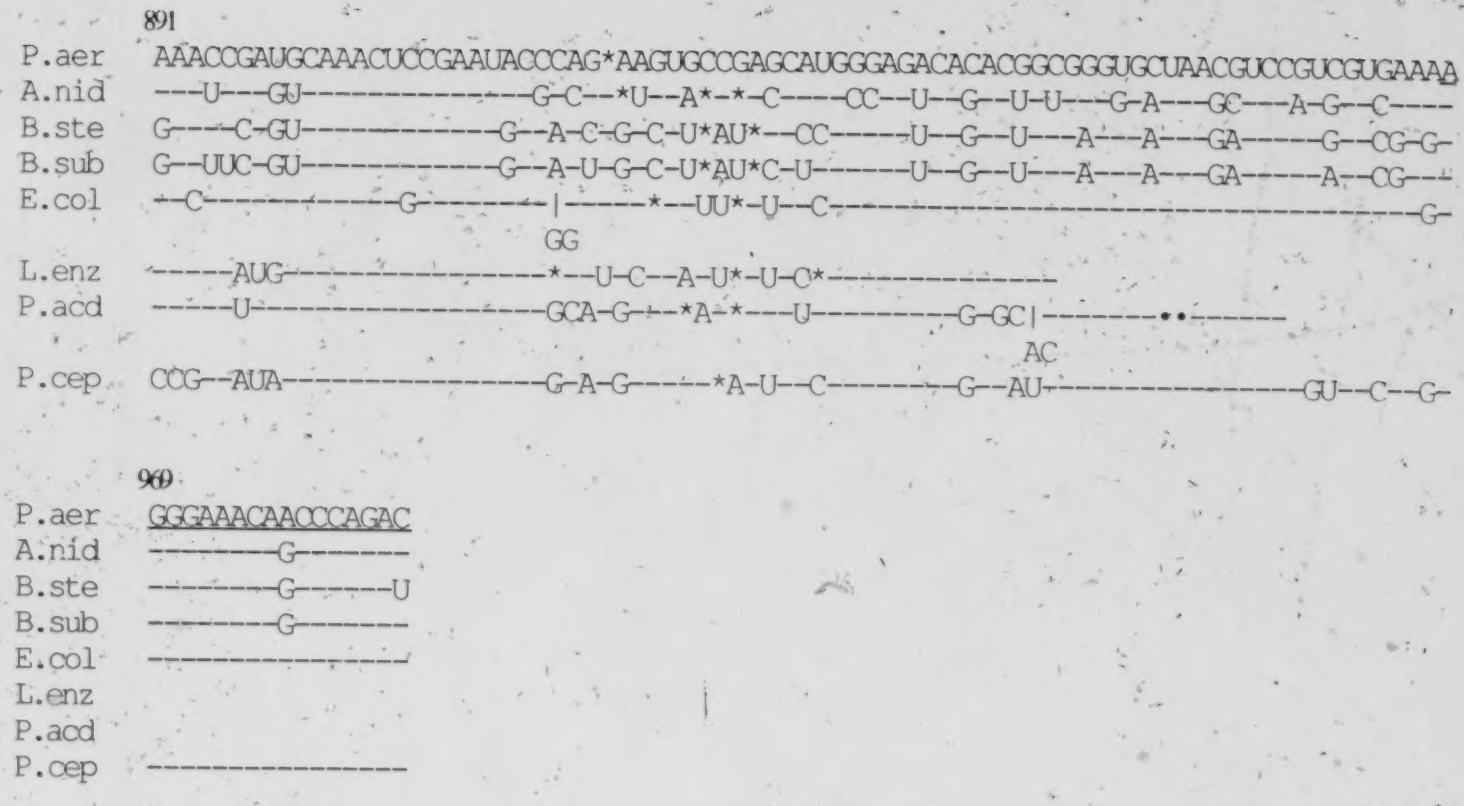


Figure 16 (continued).

Figure 17. Partial sequence alignments of selected eubacterial species with primer #3. Nucleotide positions 1571-1824 are included here. All symbols are as described in the legend to Fig. 9 and abbreviations as in Fig. 15. _____ : the hybridizing site of primer #8 (1580-1596). The sequences from positions 1700-1741 (helix #63) has been displayed but not aligned due to its variable nature.

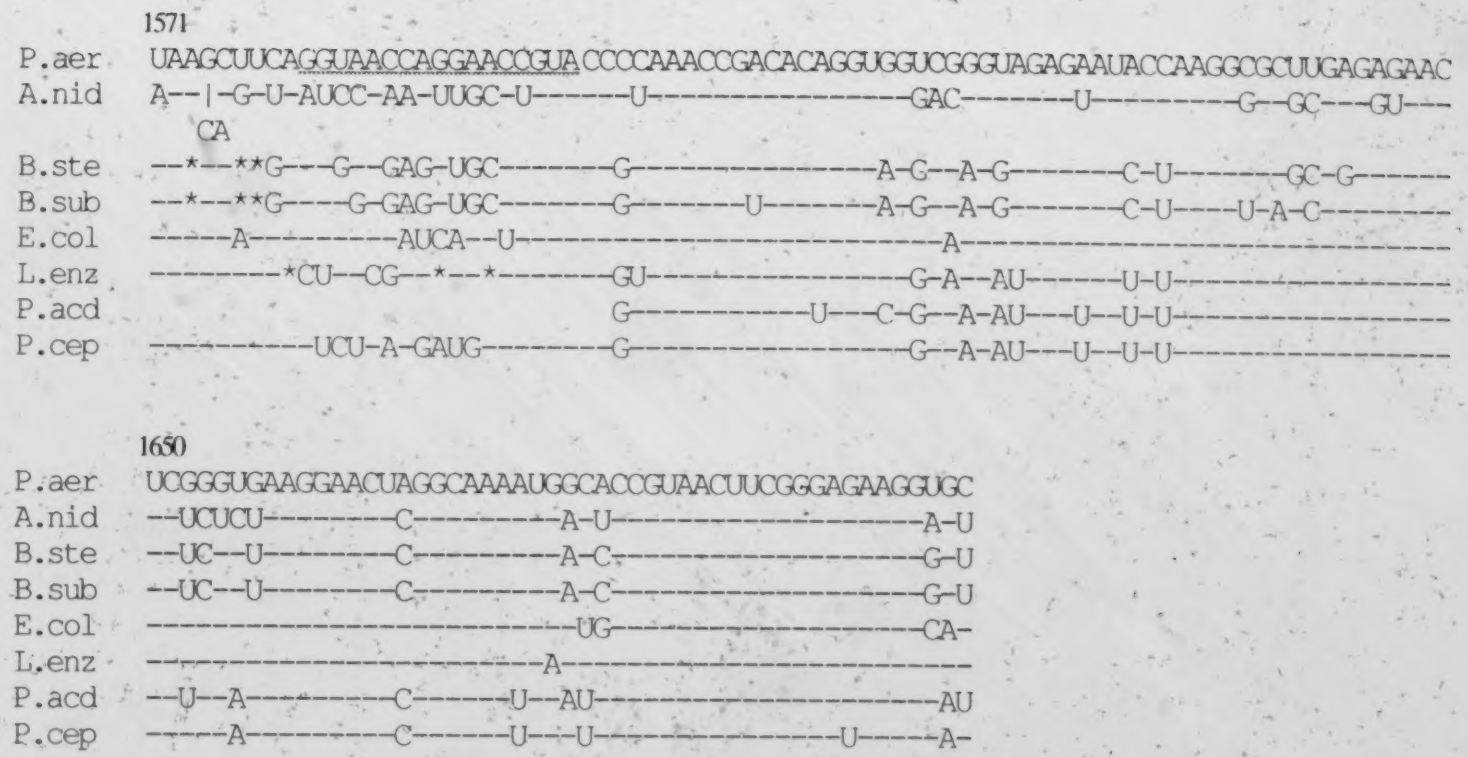

1571
P.aer UAAGCUUCAGGUAACCAGGAACCGUA CCCCAAACCGACACAGGUGGUCGGGUAGAGAAUACCAAGGCGCUUGAGAGAAC
A.nid A--|-G-U-AUCC-AA-UUGC-U------U----------GAC-------U---------G--GC---GU---
CA
B.ste --*--**G---G--GAG-UGC-------G----------A-G--A-G-------C-U-------GC-G------
B.sub --*--**G----G-GAG-UGC-------G-------U-------A-G--A-G-------C-U----U-A-C--------
E.col -----A---------AUCA--U----------A----------
L.enz --------*CU--CG--*--*-------GU----------G-A--AU------U-U----------
P.acd G----------U---C-G--A-AU---U--U-U----------
P.cep ----------UCU-A-GAUG--------G----------G--A-AU---U--U-U----------
1650
P.aer UCGGGUGAAGGAACUAGGCAAAAUGGCACCGUAACUUCGGGAGAAGGUGC
A.nid --UCUCU---------C---------A-U----------A-U
B.ste --UC--U---------C---------A-C----------G-U
B.sub --UC--U---------C---------A-C----------G-U
E.col ----------UG----------CA-
L.enz ----------A----------
P.acd --U--A----------C------U--AU----------AU
P.cep -----A---------C------U---U----------U-----A-

Figure 17 (continued).

```
1700
P.aer  GCCGGCUAGGGUGAAGGAUUUACUCCGUAAGCUCUGGCUGGU
A.nid  GCCCACCUAAGACGUGGGU
B.ste  GCUCGUUUGGGUGAAGAGCCCGAACGAGC
B.sub  GCUCUGUUAGGGUGCAAGCCCGAGAGAGC
E.col  GCUGAUAUGUAGGUGAGGUCCCUCGCGGAUGGAGCUGAAAUCAGU
L.enz  GCCCUCCAUGGUUaAUaGCUaAGGGGGGu
P.acd  GCGGCAAGUAGGUGAACUUGAACAAAGGGAGCCAAAAGCGGu
P.cep  GCCCUUGUAGCUUUGACUGGCCUGCGCCCAGGAGGGUGAAGGGGU

1742
P.aer  CGAAGAUACCAGGCCGCUGCGACUGUUUAUUAAAAACACAGCACUCUGCAAACACGAAAGUGGACGUAUACGGUGUGAC
A.nid  --C--UG-AG-----CAG-----------CC-----------GU---C--U--GU--U---AC--U-----G---GC----
B.ste  --C---C-AA-----CAA-----------C------------GU-----------G--GU--U---GC-------------GC----
B.sub  --C--UG-AU-----CAG-----------GC-----------GU-----------G---GC--U---GC--A-------------GC----
E.col  ----------CUG-----A-----------------------G-------------------------C-------------------
L.enz  U-C•-UG-----------------------C-----------u-------------------------u-------------------
P.acd  U-C•A-A-AUC--UG-----------------A-----•-•-----------•-------------•-------------------
P.cep  U-C-AUA-A-U--UG-----------------A-----------------------------------------------------
```

Figure 17 (continued).

```
1821
P.aer  GCCUGCCCGGUGCCGGAAGG
A.nid  ------A----------
B.ste  A-----------U-----
B.sub  ------------U-----
P.cep  --------------------
E.col  --------------------
L.enz
P.acd
P.cep  ----------------A---
```

Figure.18. Partial sequence alignments of selected eubacterial species with primer #7. This alignment ranges from 2164-2382. All symbols are as described in the legend to Fig. 9 and abbreviations as in Fig. 15.

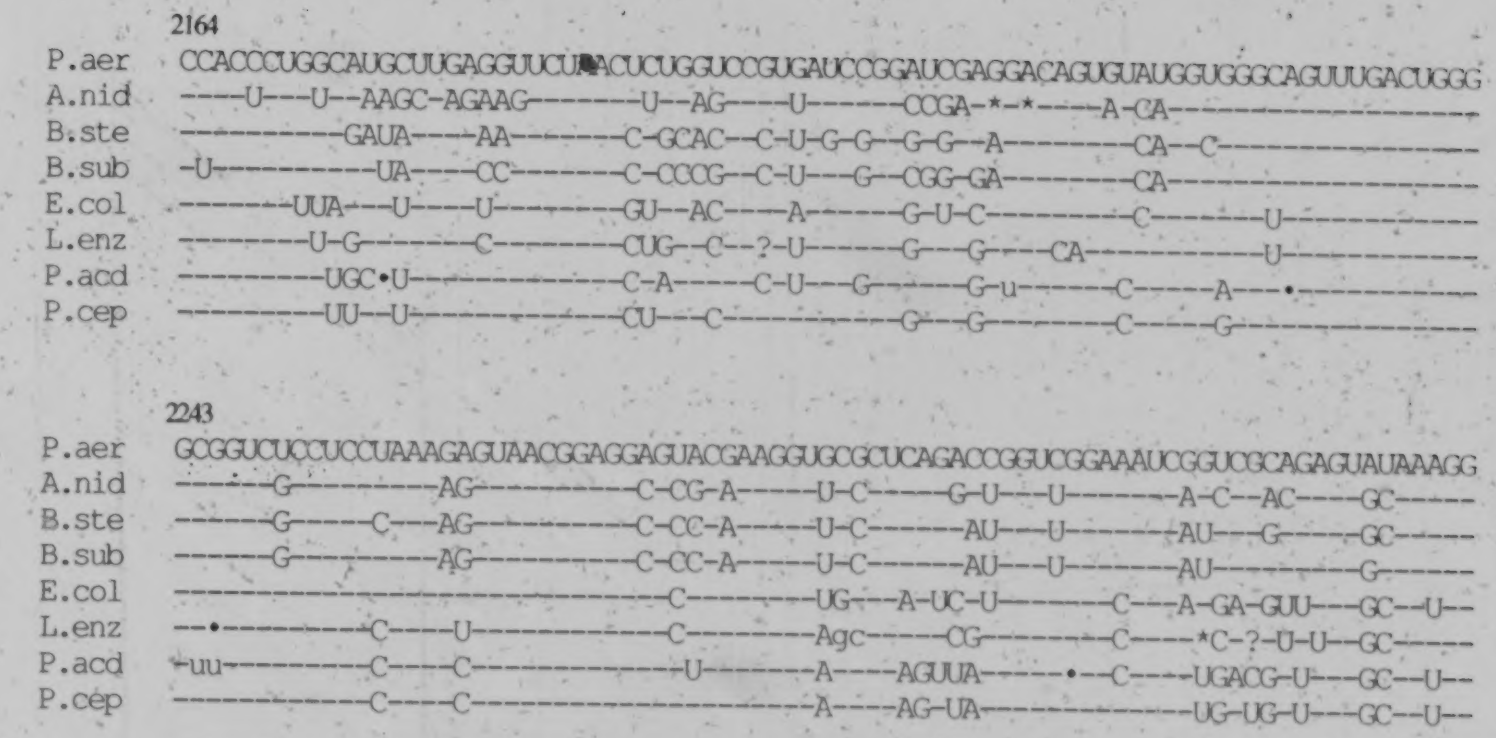
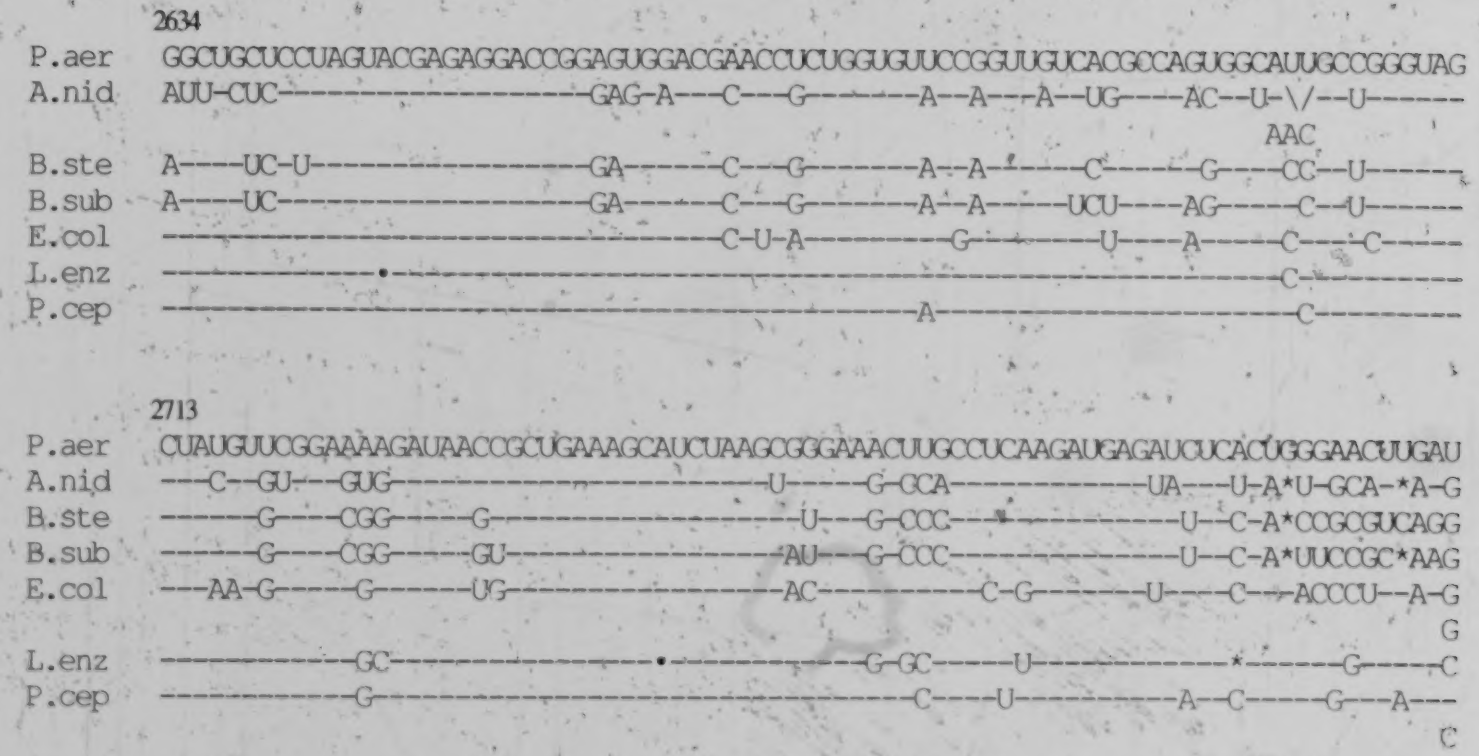

2164
P.aer CCACCCUGGCAUGCUUGAGGUUCUAACUCUGGUCCGUGAUCCGGAUCGAGGACAGUGUAUGGUGGGCAGUUUGACUGGG
A.nid ----U---U--AAGC-AGAAG-------U--AG----U------CCGA-*-*----A-CA----------
B.ste ----------GAUA----AA-------C-GCAC--C-U-G-G--G-G--A--------CA--C----------
B.sub -U----------UA----CC-------C-CCCG--C-U---G--CGG-GA--------CA----------
E.col -------UUA---U----U--------GU--AC----A------G-U-C---------C-------U----------
L.enz --------U-G-------C--------CUG--C--?-U------G---G---CA----------U----------
P.acd ---------UGC•U-------------C-A-----C-U---G------G-u------C-----A---•----------
P.cep ---------UU--U-------------CU---C-----------G---G--------C-----G----------
2243
P.aer GCGGUCUCCUCCUAAAGAGUAACGGAGGAGUACGAAGGUGCGCUCAGACCGGUCGGAAAUCGGUCGCAGAGUAUAAAGG
A.nid ----·--G---------AG----------C-CG-A-----U-C-----G-U---U-------A-C--AC----GC-----
B.ste ------G-----C---AG----------C-CC-A-----U-C------AU---U-------AU---G-----GC-----
B.sub ------G----------AG----------C-CC-A-----U-C------AU---U-------AU---------G------
E.col ------------------------------C--------UG---A-UC-U-------C---A-GA-GUU---GC--U--
L.enz --•---------C----U-----------C--------Agc-----CG--------C-----*C-?-U-U--GC-----
P.acd -uu---------C----C-------------U-------A----AGUUA------•--C----UGACG-U---GC--U--
P.cep ------------C----C--------------------A----AG-UA-------------UG-UG-U---GC--U--

Figure 18 (continued).

```
2322
P.aer  CAAAAGCGCGCUUGACUGCCGAGACAGACACGUCGAG.c ''' JACGAAAGUAGGUCUUAGUGAUCCGGUGGUUCUGUAUGG
A.nid  --U---G-A-----------A-----CU---A-----------------G-----------C---C-----------AC-----------AG---
B.ste  --C---G-A-----------A-----G---G-----------------G-----------C---G-----------C-C----------
B.sub  --C---G-A-----------A-----CU---A-----------------G-----------C---G-----------C-C----------
E.col  --U---CA----------AGU---GGCG---------------------G-----------C---A-----------A----------
L.enz  --U---U-----------•-----UC---GGA---A-----------•-----------
P.acd  --U---A---A-------------U---A-----------------A-G-----------C---A-A---------
P.cep  --U---U---A-------------C---A-----------------G-----------C---A-----------
```

Fig 19. Partial sequence alignments of selected eubacterial species with primer #5. Sequences in this alignment range from 2634- 2874. All symbols are as described in the legend to Fig. 9 and abbreviations as in Fig. 15.


2634
P.aer GGCUGCUCCUAGUACGAGAGGACCGGAGUGGACGAACCUCUGGUGUUCCGGUUGUCACGCCAGUGGCAUUGCCGGGUAG
A.nid AUU-CUC----------GAG-A---C---G-------A--A---A--UG----AC--U-\/--U------
AAC
B.ste A----UC-U----------GA------C---G-------A--A-----C------G----CC--U------
B.sub A----UC----------GA------C---G-------A--A-----UCU----AG-----C--U------
E.col ----------C-U-A---------G---------U----A-----C----C------
L.enz ----------•----------C---------
P.cep ----------A----------C---------
2713
P.aer CUAUGUUCGGAAAAGAUAACCGCUGAAAGCAUCUAAGCGGGAAACUUGCCUCAAGAUGAGAUCUCACUGGGAACUUGAU
A.nid ---C--GU---GUG----------U-----G-GCA----------UA---U-A*U-GCA-*A-G
B.ste ------G----CGG-----G----------U---G-CCC----*----------U--C-A*CCGCGUCAGG
B.sub ------G----CGG-----GU----------AU---G-CCC----------U--C-A*UUCCGC*AAG
E.col ---AA-G-----G------UG----------AC----------C-G-------U----C---ACCCU--A-G
G
L.enz ----------GC----------•----------G-GC-----U----------*------G-----C
P.cep ----------G----------C----U----------A--C-----G---A---
C

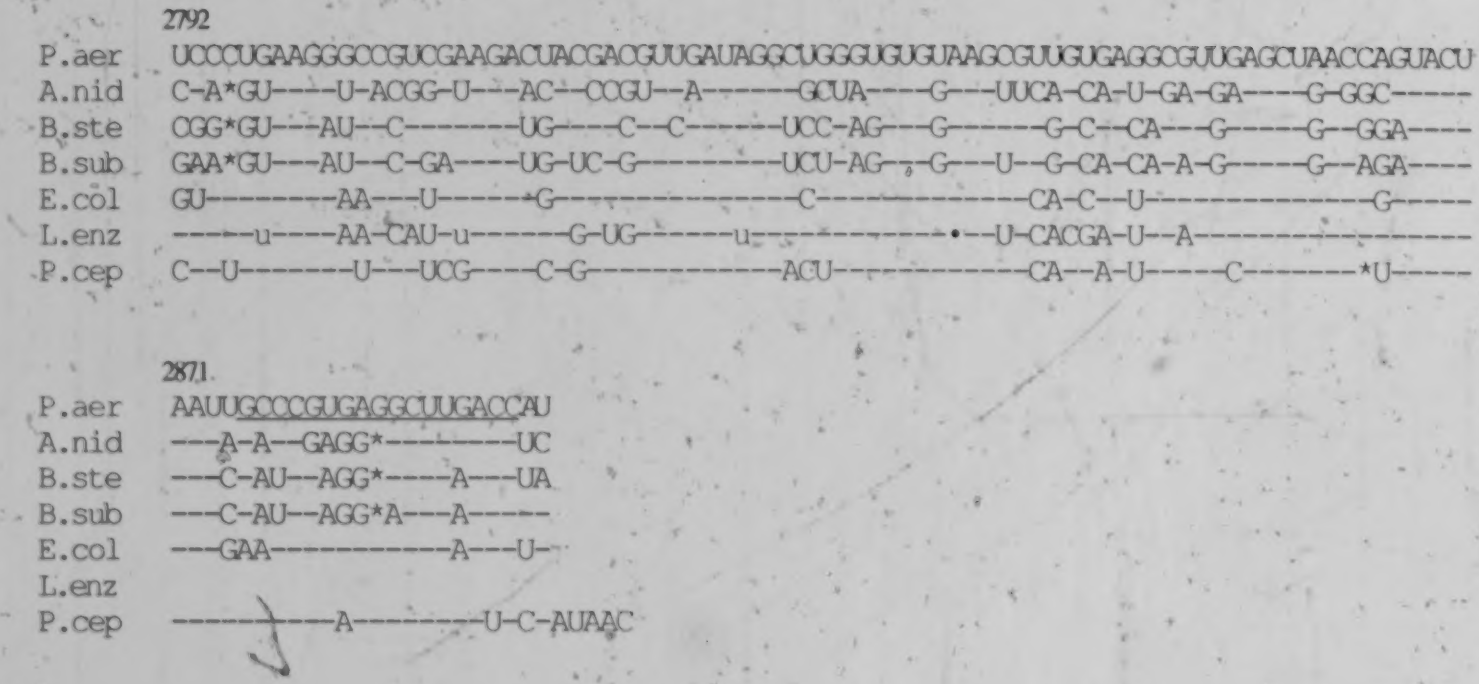


Figure 19 (continued).

TABLE 4. Pairwise comparison of selected eubacterial species.

Percentage dissimilarities (lower left) and K_{nuc} values (upper right) for each pair of the selected eubacterial spp as are displayed. These values were calculated from the data in figs. 15-19. The abbreviations used are the same as described in the legend to fig. 15.

	P.aer	A.nid	B.ste	B.sub	E.col	L.enz	P.acd	P.cep
P.aer		0.36	0.3	0.26	0.17	0.17	0.2	0.18
A.nid	28.6		0.28	0.27	0.39	0.39	0.41	0.41
B.ste	24.7	23.3		0.12	0.33	0.33	0.29	0.32
B.sub	21.7	22.8	10.7		0.34	0.36	0.30	0.34
E.col	15.3	30.4	26.8	27.1		0.23	0.21	0.24
L.enz	15.5	30.5	26.9	28.3	19.8		0.22	0.21
P.acd	17.3	31.4	24.4	25.0	18.4	18.9		0.13
P.cep	15.9	31.3	25.9	27.0	20.2	18.4	11.6	

Figure 20. Unrooted phylogenetic tree showing the relationships among the authentic *Pseudomonas* species. This tree was derived from the data presented in Table 3. The scale bar represents a K_{nuc} value of 0.005. Altogether there are 4 clusters: one with 6 members, one with 4 members and two with single members. the branch lengths between each of these represents the average distance between them. The branch lengths showing distances between the species within are cluster are not valid when considering indviduals from different clusters.

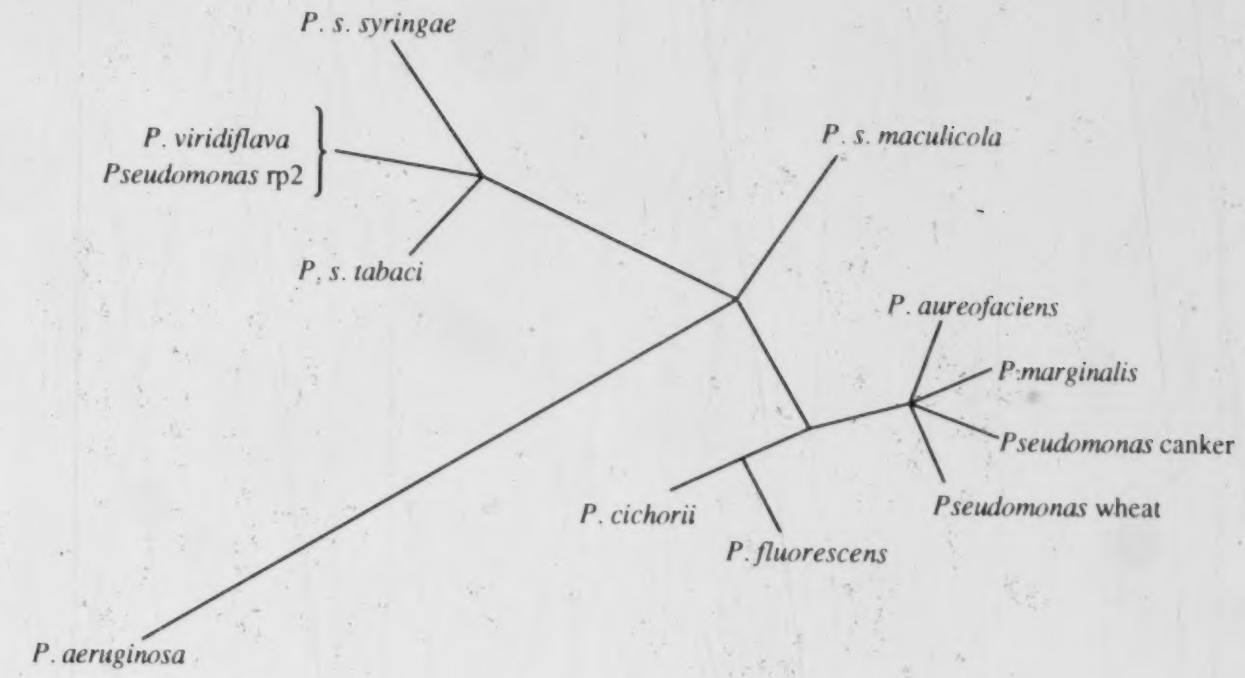
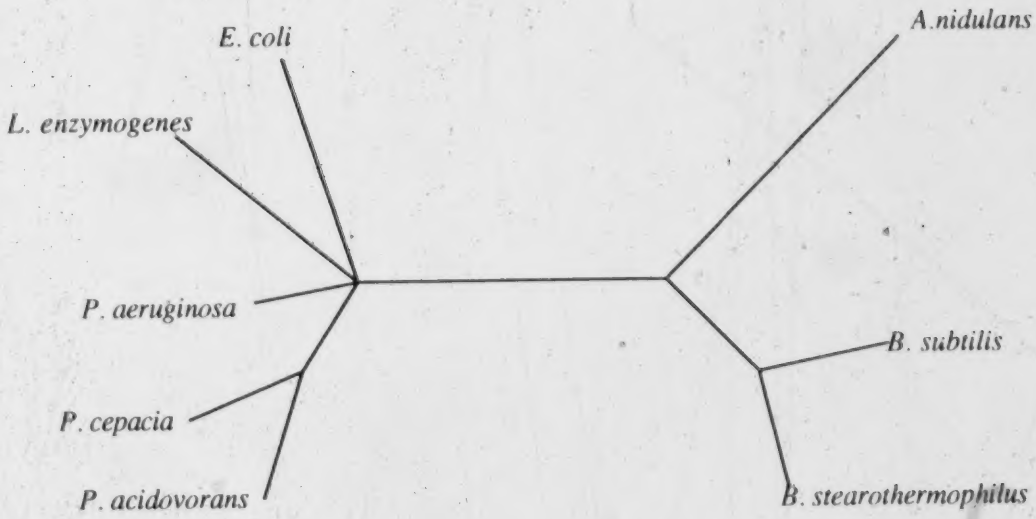

P. s. syringae
P. viridiflava
Pseudomonas rp2
P. s. tabaci
P. s. maculicola
P. aureofaciens
P. marginalis
Pseudomonas canker
Pseudomonas wheat
P. cichorii
P. fluorescens
P. aeruginosa
K nuc =0.005

Fig. 21 Unrooted phylogenetic tree showing the relationships among selected eubacterial species. This tree was derived from the data presented in Table 4. The scale bar represents a K_{nuc} value of 0.05. The distance between any two organisms is the total length of the branches linking them.


E. coli
A.nidulans
L. enzymogenes
P. aeruginosa
B. subtilis
P. cepacia
P. acidovorans
B. stearothermophilus

$K_{nuc} = 0.05$

DISCUSSION

RFLP analysis

The autoradiograms from the RFLP analysis clearly demonstrate the presence of multiple copies of the rRNA genes in the *Pseudomonas* genome. Hartmann *et al.*, (1986) have reported the presence of 4 ribosomal RNA operons in *P. aeruginosa*, but both *Eco*RI and *Eco*RV demonstrate that there are at least 5 or 6 copies of the rRNA operon per genome in all the organisms in this study. This is indicated by the presence of 5 or 6 equally intense bands in all the lanes (see Figs. 5 and 6). The different RFLP patterns are a reflection of differences in the flanking sequences of the rRNA genes. Smaller fragments which are seen in the *Hin*dIII digests of the bacterial genome are due to sites within the rRNA genes that are recognized by the restriction enzyme. When such sites are present a larger number of fragments can bee seen. A single site within the gene cluster would generate twice as many bands as there are genes and these would be less intense. The presence of 2 internal sites would yield one very strongly hybridizing band of relatively small size, and twice as many weakler bands. The rRNA genes of *L. enzymogenes* clearly reflect the presence of internal cleavage sites for all three of the enzymes tested.

In this study this technique was used as a preliminary tool to investigate possible relationships among the various *Pseudomonas* species. A tentative identification of various field isolates was possible by comparing their RFLP patterns to a battery of those from known organisms. It is a good idea to use more than one enzyme in such an analysis, to verify the results and decrease the possibilities of misinterpretations regarding identity. For example, in our analysis the hybridization patterns of the *Eco*RI digests of *P. viridiflava* and *Pseudomonas* rp2 are identical (Fig 1 lanes 10 and 11) but the *Eco*RV digests are not (Fig 2 lanes 7 and 11). The use of

more than one probe may increase the sensitivity of this test. However rRNA probes are most useful because of the multiplicity of their genes.

In general, the greater the number of common bands between two species the more closely they are related. Neither *Pseudomonas* wheat nor *Pseudomonas* canker are identical to any of the known species in this study, but both have at least 2 or 3 bands in common with other species namely *P. aureofaciens*, *P. fluorescens*, and *P. marginalis*. It is important to stress that these RFLP patterns do not guarantee homologies between the sequences of the rRNA genes themselves, but rather reflect the arrangement of these genes in the bacterial genome. In addition to providing clues about the identities of our field isolates, this study served to re-emphasize the fact that the pathovars of *P. syringae* are a heterogeneous group of bacteria with as much resemblance to each other as to other closely related plant pathogenic species.

Technical aspects of the sequencing analyses.

Sequence analyses were conducted using the methodology described by Lane *et al.* (1985) for the rapid determination of rRNA sequences using reverse transcriptase-catalyzed primer extension. The advantage of using this approach was that large amounts of template could be isolated easily and rapidly and used directly without the need for cloning into a vector. Sequences in the range of 220-250 nt in length were obtained from each primer yielding a total of 1220-1400 nt per organism.

The primers used in these experiments were chosen by comparing the known 23S rRNA sequences of *E. coli* and *P. aeruginosa* (Gutell and Fox, 1988). Small tracts of sequences that were conserved between these two species, and lying towards the 3' direction of variable regions of the 23S rRNA molecule were selected as primer hybridizing sites, the underlying assumption being that a region common to these two distantly related organisms would be conserved within either genus. The majority of

the organisms compared in these experiments were closely related *Pseudomonas* species (aerobic, gram negative rods), and hence the gram negative sequences appeared to be the logical choice for selecting primers. The locations and sequences of these sites are shown in Fig.7 and Table 2. The priming sites were chosen such that all six structural domains were included in the analysis.

As stated earlier, two of the six oligonucleotides originally synthesized (#2 and #4) failed to serve as suitable primers for extension. The reason for this in the case of primer #2 was discovered on reexamining its sequence. The oligonucleotide was synthesized in the incorrect orientation and hence it was not complementary to the proposed hybridizing site. The failure of primer #4 could probably be attributed to the location of its hybridizing site in the RNA molecule. The 23S rRNA is a complex molecule with extensive secondary structure. In a typical *in vitro* sequencing reaction. the template and primer are first heated at 90°C and then allowed to cool to the temperature of the elongation reaction (50°C in our experiments). This is to ensure the complete denaturation of the template at 90°C, so that the primer can anneal to it properly and initiate DNA synthesis at 50°C. However some stable double helical and stem-loop (hairpin) structures found in RNA molecules are capable of reforming almost instantaneously on cooling to 50°C.

An examination of the hybridizing sites of the primers revealed that #4 immediately preceded a region of the 23S rRNA molecule with extremely stable double helical structures. These structures would probably reanneal at 50°C, rendering themselves inaccessible to reverse transcriptase. Therefore, chain elongation could not take place. In contrast , the hybridizing sites of the successful primers are located within stable stem–loop or double helical structures of the RNA molecule. When these primers anneal to their complementary sites, they prevent the RNA molecule from folding back on itself. This opens up the molecule and makes it available as a template for DNA synthesis. Using this rationale, it appeared as though primer #2 would

probably have been ineffective even in the correct orientation. Therefore, primers #7 and #8 were synthesized as replacements for #4 and #2 respectively. Their sequences extended into the base paired region adjacent to the original priming sites (see Fig. 7) and they proved effective as primers.

Additional difficulties were encountered when dealing with *L. enzymogenes* and *P. acidovorans*. Sequence determination with primer #8 was not possible for either of these these two organisms, while primer 5 failed to initiate synthesis in *P. acidovorans*. The simplest explanation would appear to be a lack of complementarity between the primer and template which prevented them from annealing to one another. Fig 17 which shows the sequence of *L. enzymogenes* obtained from primer #3 extending into the region of primer #8, offers evidence to support this explanation. Although direct proof of non-complementarity is unavailable for *P. acidovorans*, it does seem plausible on the basis of the available data. Comparisons between the 23S rRNA sequences of *P. aeruginosa* and *P. cepacia* (Höpfl *et al.*, 1989) show dissimilarities in the regions of these 2 primers (see figs 17 and 19). While *P. cepacia* and *P. acidovorans* are by no means identical, they are more closely related to one another than to *P. aeruginosa*. This fact was first indicated by DeVos and DeLey (1983) based on rRNA hybridization data. Furthermore, the data obtained in this study, from the analysis of selected regions of the molecule (domains I, II, IV and V: see Figs. 15-18) also give evidence for this relationship.

Fig 8 is a typical example of the type of sequence data obtained in these experiments. Bands of equal intensity are observed in all 4 lanes at the position indicated by the bold arrowhead and the nucleotide at that position cannot be identified. These types of anomalies have been reported previously (Lane *et al.*, 1985), and are a result of terminations which are not due to ddNTP incorporations. An examination of the sequence alignments (Figs. 9-22), reveals that these non-specific terminations

usually occurred at equivalent sites in RNAs from different organisms, indicating that the problem could be related to the structure of the template molecule at those positions.

One possible explanation is that a proportion of the RNA molecules are nicked at various sites, producing a mixed population of template molecules of different lengths. Owing to the extreme stability of RNases, this is an extremely likely event. When the reverse transcriptase encounters these breaks, it terminates synthesis regardless of the ddNTP in the reaction mixture, giving rise to bands in all the lanes. Chain elongation proceeds normally in the longer fragments and intact molecules which allows the reading of sequence beyond these points. It is reasonable to assume that the sensitivity of various sites to RNase is directly related to the secondary structure of the 23S rRNA; helical regions with extensive base–pairing would be more resistant, while exposed, unpaired regions of the molecule would be more susceptible to enzymatic cleavage. The locations of the ambiguously identifiable sites (see sequence alignments) appear to support this line of reasoning. An alternate possibility, suggested by Lane *et al.* (1985), is that some of the non specific terminations might arise from modified residues in the template, or as a result of tertiary structure interactions which may cause the reverse transcriptase to pause before resuming synthesis. It is conceivable that all of these factors might play a role in generating the ambiguities.

Sequence alignments

The sequence alignments are presented in order of the occurrence of the primers in the rRNA molecule from the 5' end of the molecule. Alignments were conducted by eye to ensure that analogous structures from different rRNAs were compared. This was especially important when considering the more distant eubacterial species, where corresponding helices or loops vary in length in different organisms. The rRNA sequences of the authentic *Pseudomonas* species could be aligned more easily with

respect to their primary structure. The following discussion deals with the characteristics of the sequenced regions within each domain. Specific helices are referred to by numbers according to the scheme of Leffers *et al.* (1988).

Primer #1 which lies in domain I gave information about the region extending from helix 9 to 22 (refer to Figs 9, 15 and 22). The primer hybridizing site lies within helices 14 and 22 and is well conserved among the gram negative organisms *P. aeruginosa, E. coli,* and *P. cepacia* (Fig 15) and probably also *P. acidovorans* and *L. enzymogenes*. Maximum variability was seen in the region of helices 9,16 and 18. Helix 9 is absent from the *A. nidulans* 23S rRNA and varies considerably both in size and primary sequence among the other organisms. Within the genus *Pseudomonas* this helix shows differences in its primary sequence but appears constant with respect to size. The portion of the molecule including the helices 16 and 18 appears to evolving at a fairly rapid rate. The helices themselves are non-alignable among different genera and show sequence variation among the different species of the same genus. Helix 16 contains insertions of up to 40 nt in the *Bacillus* spp. Also in this region are 2 additional features; helix 15 (identical to 16) which is seen in the archaebacteria , and helix 17 which is unique to the archaebacterium *Desulfurococcus mobilis* (Leffers *et al.*, 1988).

Sequences from domain II were determined using primer #6 (figs 10, 16 and 23). This region which includes helices 32-41, seems very well conserved within the genus *Pseudomonas*. A maximum of 8 differences from the sequence of *P. aeruginosa* is observable out of a total of 222 nucleotide positions and the differences between the other species are even fewer. All of the secondary structure elements in this region seem well conserved with no significant deletions or insertions although there is more variation in the primary sequence among the more distant eubacteria.

Sequence data from domain III were unobtainable for the organisms *P. acidovorans* and *L. enzymogenes*. As explained before, this was probably due to a

Figure 22. Secondary structure model of domain I of the 23S rRNA of *P. aeruginosa* (enlarged from Fig. 7). The hybridizing site of primer #1 is outlined, specific helices are indicated by arabic numerals, and the nucleotide positions indicated by a number with a pointer.

Figure 22: Secondary structure model of domain I of the 23S rRNA of *P. aeruginosa*

reproduced from: R.R. Gutell and G.E. Fox (1988). Compilation of large ribosomal subunit RNA sequences presented in a structural format. Nucleic Acids Research 16 (sequence supplement) 175-270.

has been removed pending receipt of copyright permission.

Figure 23. Secondary structure model of domain II of the 23S rRNA of *P. aeruginosa* (enlarged from Fig. 7). The hybridizing site of primer #6 is outlined, specific helices are indicated by arabic numerals, and the nucleotide positions indicated by a number with a pointer.

Figure 23: Secondary structure model of domain II of the 23S rRNA of *P. aeruginosa*

reproduced from: R.R. Gutell and G.E. Fox (1988). Compilation of large ribosomal subunit RNA sequences presented in a structural format. Nucleic Acids Research 16 (sequence supplement) 175-270.

has been removed pending receipt of copyright permission.

lack in complementarity between primer #8 and its hybridizing site (helix 54) in these 2 organisms. Fig 17 shows this region to be quite variable amongst the other eubacterial species. The sequence alignment in Fig 12 suggests that this is not the case among the authentic *Pseudomonas* species, as the sequences which did extend into that region like those of *P. s. syringae* and *Pseudomonas* canker appear conserved. Furthermore, the ability of primer #8 to initiate synthesis implies its complementarity to the template sequence. Helices 49-59 were sequenced using this primer. Maximum variability is observed in helices 57 and 59 (Fig 24).

Primer #3 allowed the determination of the sequences of helices 60-68 (Figs 12. 17 and 24) in domain IV. Helix 63 shows a very high degree of variability in size in the eubacteria examined and hence cannot be aligned. Once again the helix is of a constant size within the genus *Pseudomonas* but shows differences in the primary sequence. The primer hybridizing site itself occurs in the helix 68 which is highly conserved. 7 out of the 16 nucleotides are invariant in all of the large subunit rRNAs sequenced to date (Höpfl *et al.*, 1989).

The region sequenced in domain V is almost invariant in the genus *Pseudomonas*, and all secondary structure features seem well conserved even in the more distant eubacteria (Figs 13, 18 and 26). A notable feature in this domain is the universal occurrence of tertiary interaction between two triplet sequences. The first sequence lies in the region between helices 83 and 84 (positions 2318-2320 in *P. aeruginosa*) and the second between 87 and 88 (2375-2377). The triplet sequences themselves are not invariant but a change in one position is always accompanied by a compensating change in the second triplet. In *P. aeruginosa* the triplet AAG is linked to CUU, in *E. coli* the triplets are AUG and CAU. The triplet pair is constant in all of the authentic *Pseudomonas* spp we examined, as well as in the other eubacteria with the exception of *E. coli*, *P. cepacia*, and *P. acidovorans*.

Figure 24. Secondary structure model of domain III of the 23S rRNA of *P. aeruginosa* (enlarged from Fig. 7). The hybridizing site of primer #8 is outlined, specific helices are indicated by arabic numerals, and the nucleotide positions indicated by a number with a pointer.

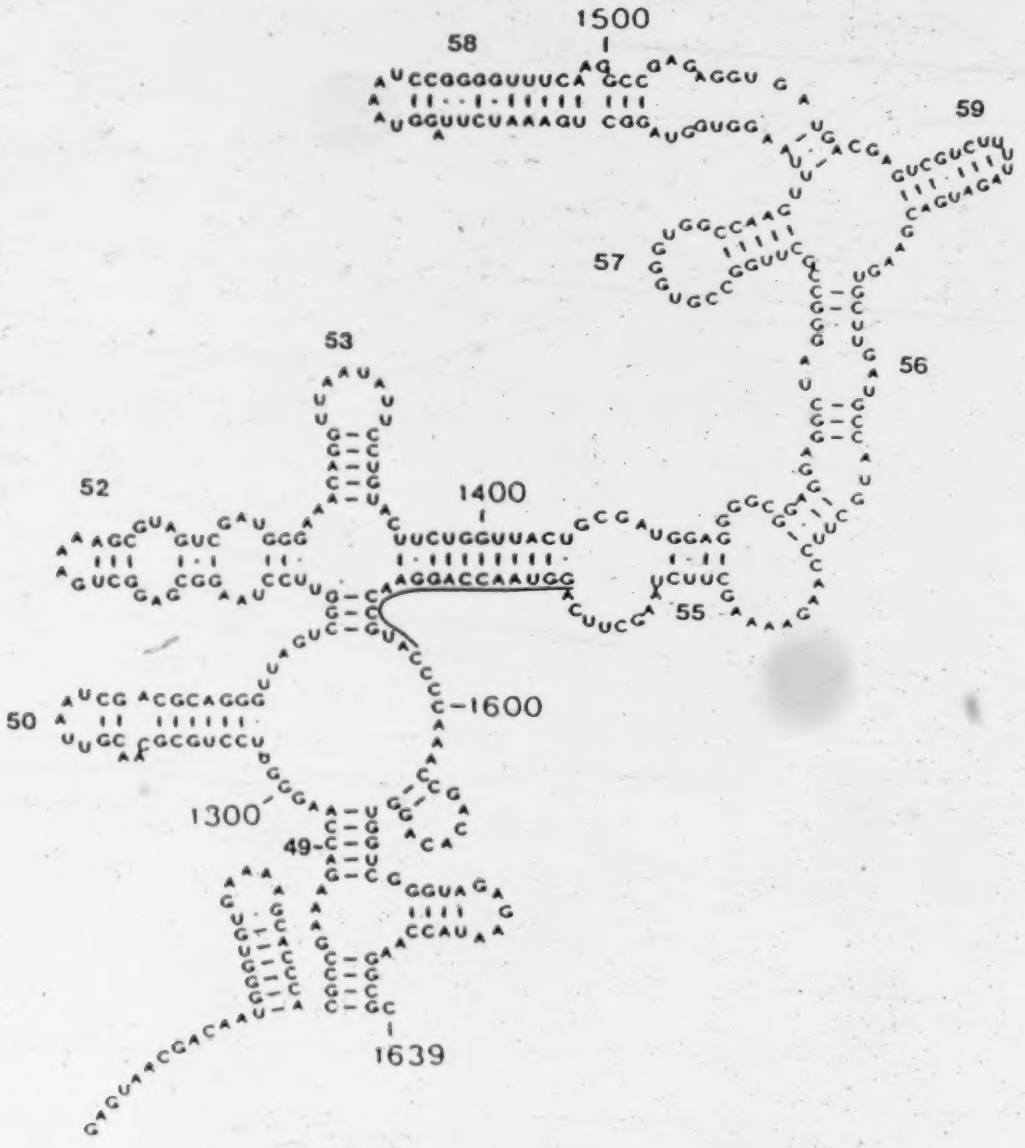

1500
58
59
57
53
56
52
1400
55
50
1600
1300
49
1639

Figure 25. Secondary structure model of domain IV of the 23S rRNA of *P. aeruginosa* (enlarged from Fig. 7). The hybridizing site of primer #3 is outlined, specific helices are indicated by arabic numerals, and the nucleotide positions indicated by a number with a pointer.

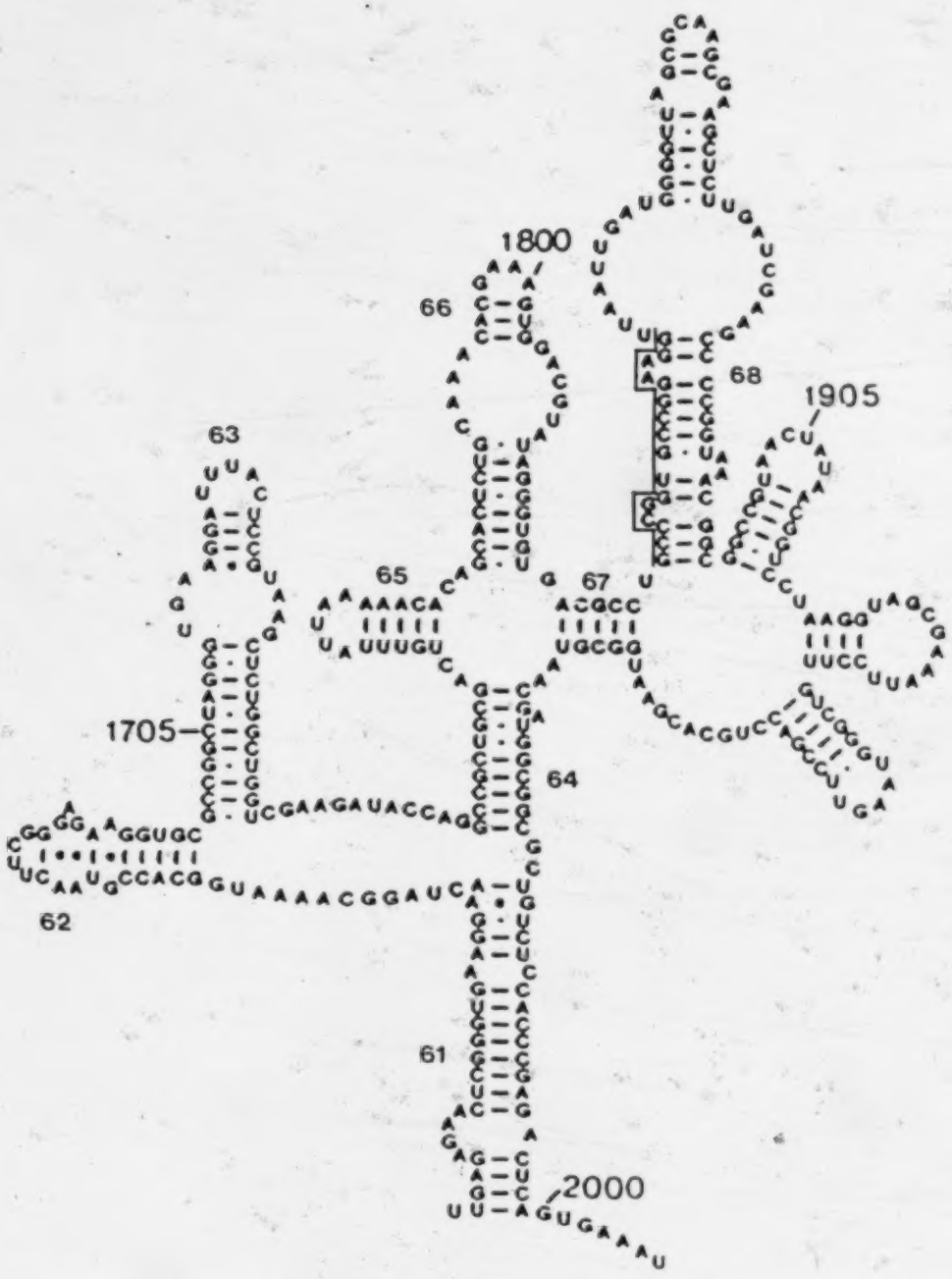
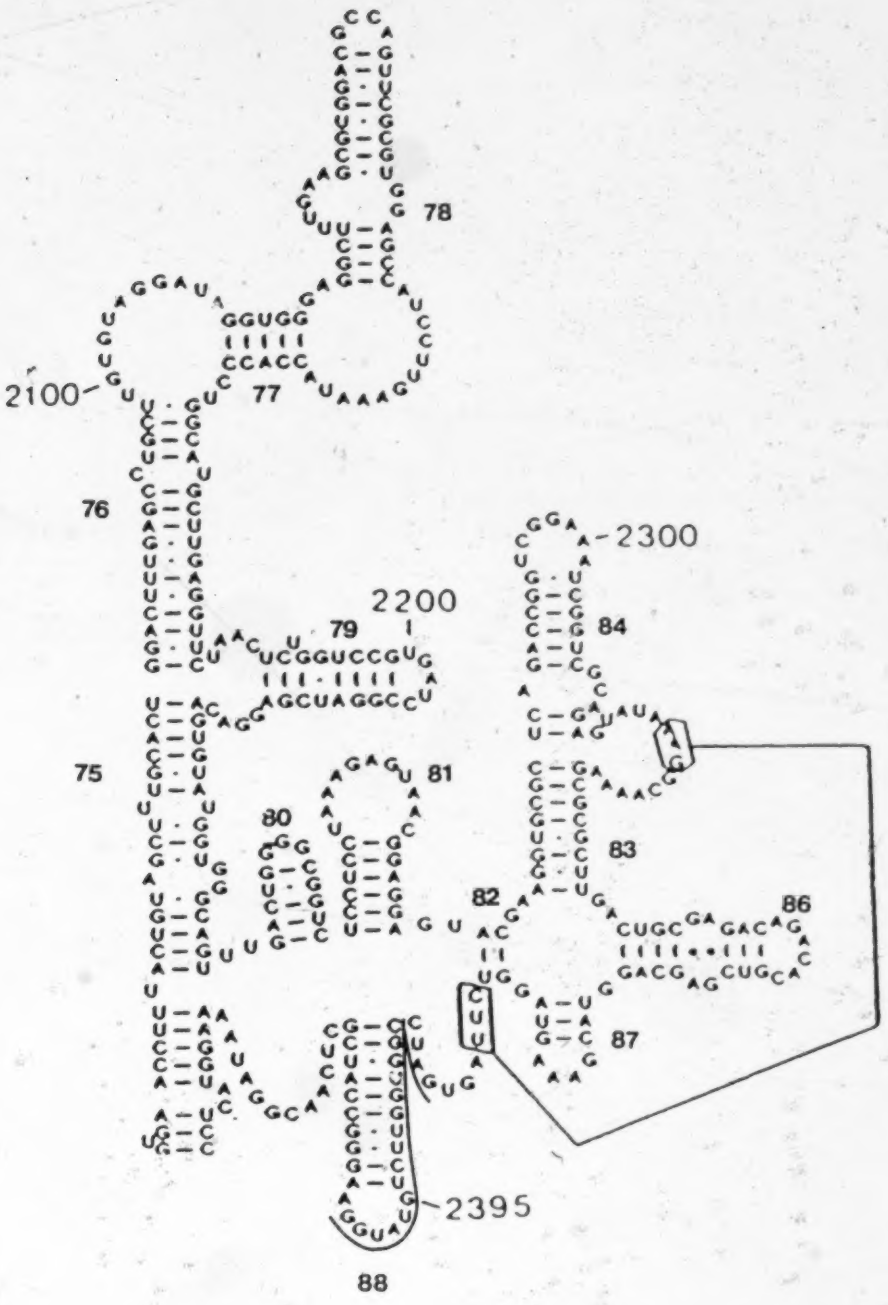

1800
1905
1705
2000
62
63
64
65
66
67
68
61

Figure 26. Secondary structure model of domain V of the 23S rRNA of *P. aeruginosa* (enlarged from Fig. 7). The hybridizing site of primer #7 is outlined, specific helices are indicated by arabic numerals, and the nucleotide positions indicated by a number with a pointer.. The triplet sequences involved in the tertiary interaction are indicated.

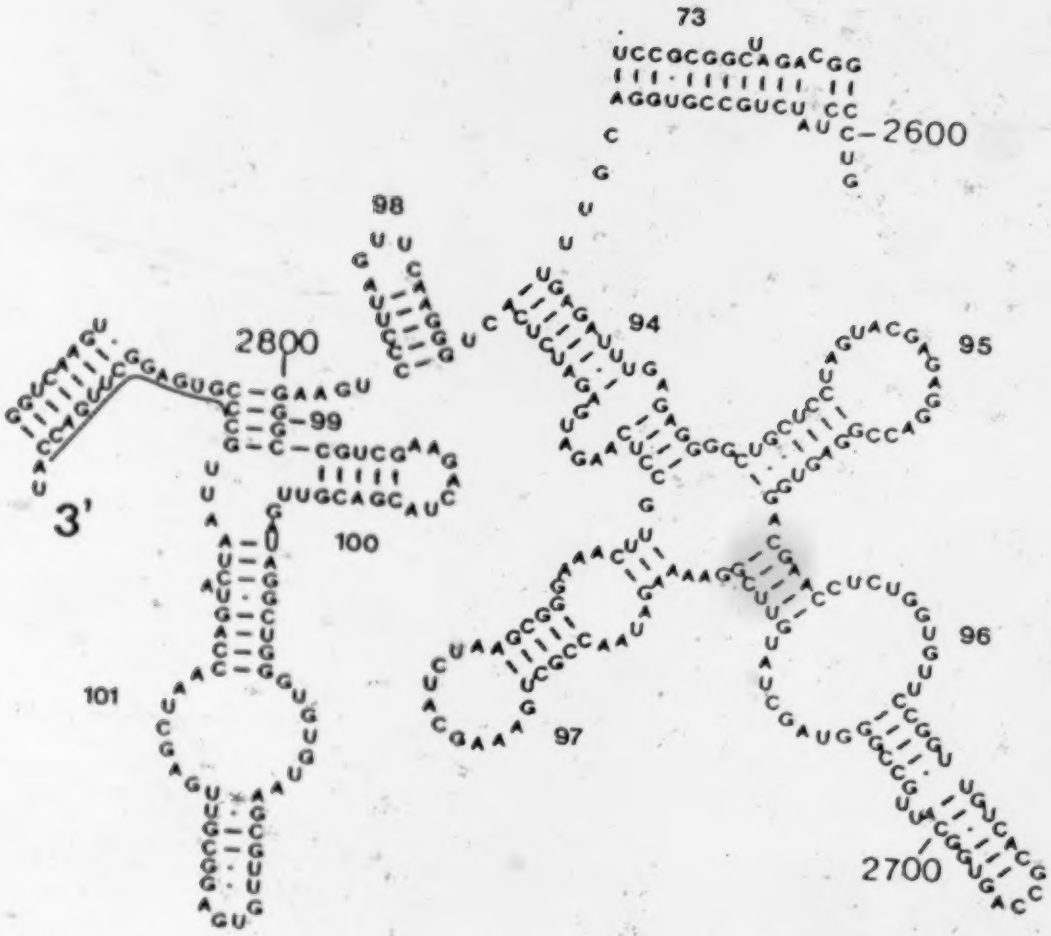

2100
2200
2300
2395
75
76
77
78
79
80
81
82
83
84
86
87
88

The 3'-end of the molecule (domain VI) was determined by primer #5 (figs 14, 19 and 27). Helix 98 is the most variable region in this domain. Two stretches with a high degree of conservation can be observed in this region; the loop region in the stem-loop structure of helix 95 (positions 2544-2655), and a portion of the helix 97 which includes nucleotides 2725 to 2760.

An examination of the entire 23S rRNA molecule shows a distinct pattern in the relative conservation and variability among different organisms. The domains with the most conserved secondary structural features are those that show the highest interspecific conservation with respect to primary sequence. Helices showing a high degree of sequence variability within a genus are virtually non-alignable between different organisms. Such helices are usually the sites for major insertions or deletions as demonstrated in the specific cases of helices 9, 16, 63 and 98. They represent regions with a rapid rate of evolution and appear to be the most useful areas for comparing closely related organisms. Thus, in our analyses the primers #1, #3 and #5 yielded the most valuable information regarding intrageneric relationships. On the other hand, regions like domains II and V are more useful for evaluating relationships between more distant organisms.

Conserved sites in the 23S rRNA molecule very probably represent functionally important domains. Höpfl *et al.* (1989) have listed specific sites in the eubacterial 23S rRNA corresponding to target sites for interaction with antibiotics, tRNAs and the elongation factors (EFs) Tu and G. In general these sites were well conserved in all of the observed species. For instance, the invariant loop of helix 95 contained 4 individual positions involved in the interaction with EF-Tu and EF-G and 2 target sites for the antibiotic α-sarcin. Most of the other sites listed are present at positions which were not in the regions included in this study. Besides their functional significance it is possible that the conserved portions of the molecule are regions which interdigitate with

Figure 27. Secondary structure model of domain VI of the 23S rRNA of *P. aeruginosa* (enlarged from Fig. 7). The hybridizing site of primer #5 is outlined, specific helices are indicated by arabic numerals, and the nucleotide positions indicated by a number with a pointer.


73
2600
98
94
95
2800
3'
99
100
101
96
97
2700

the ribosomal proteins. Such interactions pose certain constraints upon the evolution of both the RNA and the protein components of the ribosome.

While the primers that were selected proved quite adequate for determining the 23S rRNA sequences of all of the *Pseudomonas* species, they were quite obviously not of universal applicability. To be more generally applicable, the primers should be chosen such that they hybridize to well conserved sequences which are part of a stable helix in the 23S rRNA molecule. In general larger primers would be less sensitive to mismatches and more capable of initiating synthesis. However, there is a potential problem in that increasing the length of a primer might also increase the number of mismatches. Also the 3' end of the primer should be paired with the maximally conserved region.

For a primer to yield valuable information, it should lie adjacent to regions showing considerable sequence variability. In the sequence analysis of Höpfl *et al.* (1989) the 3' half of the molecule (domains IV, V, and VI) shows more conservation than the 5' half. Therefore, even though a primer in the region of helix 95 would yield sequence data, it is not a good choice for phylogenetic purposes since the helices which would be determined (up to position 2400) have a large proportion of invariant nucleotides. Primers which can initiate synthesis in regions containing both highly variable helices as well as more conserved structures are probably the most broadly applicable ones. A good example of this type is #3. The primer itself spans a helical region with several invariant positions and it yields information about both moderately conserved (66, 64, 62 and 61) and highly variable helices (63). While the former would be useful in determining relationships at an intergeneric or higher level, helix 63 would be useful at the interspecific level. It is also possible to increase the utility of the single priming site by using a degenerate mixture of oligonucleotides rather than a single sequence. This allows the annealing of at least a fraction of the oligonucleotides to the template, thus enhancing the overall utility of the primer. Domain I, in particular,

is a rapidly evolving region where such a primer would be very useful. Primer #1 binds to a moderately conserved sequence in this region and its usage could be greatly increased by using a degenerate mixture of oligonucleotides rather than a single sequence.

Bacterial phylogeny

The basic branching pattern of the phylogenetic tree obtained for the eubacteria (Fig 21) corresponds to that obtained in previous studies with both 23S and 16S rRNAs (Höpfl *et al.*, 1989 and Woese *et al.*, 1987, respectively). Two hitherto uncharacterized organisms, *P. acidovorans* and *L. enzymogenes*, were included in this analysis and their taxonomic positions with respect to the other organisms were determined. As indicated in earlier studies (DeVos and DeLey, 1983), *P. acidovorans* showed maximum homology to *P. cepacia*. All of the gram negative organisms were closer to each other than to the gram positives (see Table 4). *L. enzymogenes* showed maximum homology to *P. aeruginosa*. Sequences for two of the extremely variable helices (16 and 18) in domain I (see fig 15) were alignable between these two organisms. However, helix 63 in domain IV could not be thus compared due to a deletion in the structure.

All of the authentic *Pseudomonas* species showed a high degree of similarity to one another. Thus the evolutionary tree is not very well resolved. Quite clearly, the most distant organism is the human pathogen *P. aeruginosa* which shows 5-6% difference from all of the other organisms in all of the nucleotide positions examined. All of the other organisms differ from one another by 1-3%. The non-pathogenic species *P. aureofaciens*, *P. fluorescens*, *P. marginalis*, *Pseudomonas* sp (canker), and *Pseudomonas* sp (wheat), appear to cluster within a range of 1% as do the plant pathogens *P. syringae* pv. *tabaci* and *P. syringae* pv. *syringae*. *P. viridiflava* and

Pseudomonas rp2 appear virtually identical and it would be quite reasonable to designate the latter as a variant strain of *P. viridiflava*. Both these organisms appear to be closest to the *syringae/tabaci* pair. *P. cichorii* shows some clustering with the fluorescens group while *P. syringae* pv. *maculicola* does not seem to cluster with any of the other isolates.

It is interesting to compare these results with previous studies on taxonomy. Campbell *et al.* (1986) subjected 14 *Pseudomonas* spp to a battery of about 50 nutritional and physiological tests, in an attempt to classify them. When these data were analysed by means of TAXMAP both *Pseudomonas* rp2 and *P. viridiflava* were seen to form single member clusters, with no apparent grouping to any of other organism. These two organisms are virtually identical with respect to their rRNA sequence, which may indicate that the differences in their phenotypic properties are probably due to various environmental pressures. A major flaw is thus apparent in the conventional classification techniques. Since all the characters are given equal weight in these analyses, it is quite possible that certain properties are given undue importance, while certain other important properties are omitted.

Owing to the very limited degree of dissimilarity among the *Pseudomonas* species the evolutionary tree in Fig. 20 may in fact reflect artificial relationships. This problem may be avoided by examining the patterns of conservation in the most variable regions of the rRNA molecule. Helix 16 in domain I is a good example. Considerable variation can be seen in this region (positions 275-290 and 340 -355 of *P. aeruginosa*) and the relative patterns of difference do correspond to the clustering patterns in the evolutionary tree. Another region which should help in this way is helix 63 in domain IV. In fact this region shows more variation among the different species. In addition to the evidence offered by the individual domains, the clustering patterns indicated by the RFLP analyses are in agreement with the evolutionary tree data.

In comparison to older methods such as the rRNA hybridization (DeVos and DeLey, 1983; DeVos *et al.*, 1985), rRNA sequencing gave a much higher degree of resolution at the intrageneric level. For instance, the heterogeneity of the pathovars of *P. syringae* was not indicated by rRNA hybridization. In fact the method was only useful in differentiating between individuals at an intergeneric level and did not permit predictions about the clustering patterns among the different species of a genus. In addition to their greater precision, sequence data are more advantageous due to their cumulative nature.

Conclusions

The original aim of this research project was to determine the taxonomic location of *Pseudomonas* rp2 with respect to other plant pathogenic *Pseudomonas* species. This organism showed virtually no difference from *P. viridiflava* in its rRNA sequence Also, the RFLP patterns of the two were identical when probed with rRNA. Together, these data provide strong evidence to support the designation of *Pseudomonas* rp2 as a variant strain of *P. viridiflava*. While the other two field isolates were not identified as unambiguously as was *Pseudomonas* rp2, the rRNA data placed them in close proximity to the non-pathogenic species.

In general, rRNA sequencing appears to be a relatively rapid and reliable method for inferring relationships among organisms. Due to the simplicity and speed of the technique it has become very widely used in phylogenetic analyses. The 23S rRNA is clearly more advantageous when dealing with very closely related species. In addition it can provide a higher degree of resolution of the relationships inferred by 16S rRNA analyses.

REFERENCES

Arthur, M., Arbeit, R.D., Kim, C., Beltran, P., Crowe, H., Steinbach, S., Campanelli, C., Wilson, R.A., Selander, R.K., and Goldstein, R. (1990). Restriction fragment length polymorphisms among uropathogenic *E. coli* isolates: *pap*-related sequences compared with *rrn* operons. Infect. and Immun. **58:** 471-479.

Baroin, A., Perasso, R., Qu, L., Brugerolle, G., Bachellerie, J., and Adoutte, A. (1988). Partial phylogeny of the unicellular eukaryotes based on rapid sequencing of a portion of 28S rRNA. Proc. Natl. Acad. Sci. USA. **85:** 3474-78.

Bautsch, W., Grothues, D., and Tummler, B. (1988). Genome fingerprinting of *Pseudomonas aeruginosa* by two-dimensional field inversion gel electrophoresis. FEMS Microbiol. Lett. **52:** 255-258.

Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32:** 314-331.

Bowen, T., Stackebrandt, E., Dorsch, M., and Embley, T.M. (1989). The phylogeny of *Amycolata autotrophica, Kibdelosporangium aridum and Saccharothrix australiensis*. J. Gen. Microbiol. **135:** 2529-2536.

Campbell, J.N., Cass, D., and Peteya, D. J. (1987). Colonization and penetration of intact canola seedling roots by an oppurtunistic fluorescent *Pseudomonas* spp. and the response of host tissue. Phytopathology. **77:** 1166-1173.

Campbell, J.N., Conn, K., Sorlie, L., and Cook, F.D. (1986). Inhibition of growth in canola seedlings caused by an opportunistic *Pseudomonas* spp. under laboratory and field conditions. Can. J. Microbiol. **32:** 201-207.

Caugant, D.A., Aleksic, S., Mollaret, H.H., Selander, R.K., and Kapperud, G. (1989). Clonal diversity and relationships among strains of *Yersinia entercolitica*. J. Clin. Microbiol. **27:** 2678-2683.

Cook, D., and Sequeira, L., (1988). Classification of *Pseudomonas solanacearum* by analysis of restriction fragment length polymorphisms. Phytopathology **78:** 1501. (Abstract).

Coulson, A., Sulston, J., Brenner, S., and Karn, J. (1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA. **83:** 7821-7825.

Dams, E., Hendriks, L., Van de Peer, Y., Neefs, J.-M., Smits, G., Vandenbempt, I., and De Wachter, R. (1988). Compilation of small ribosomal subunit RNA sequences. Nucl. Acids Res. **16** (seq. suppl): 87-174

DeLey, J., Segers, P., and Gillis, M. (1978). Intra and intergeneric similarities of *Chromobacterium* and *Janthinobacterium* rRNA cistrons. Int. J. Syst. Bacteriol. **28:** 154-168.

Denhardt, D.T. (1966). A membrane filter technique for detection of complementary DNA. Biochem. Biophys. Res. Commun. **23:** 641-646.

Denny, T.P., Gilmour, M.N., and Selander, R.K. (1988). Genetic diversity and relationships of two pathovars of *Pseudomonas syringae* . J. Gen. Microbiol. **134:** 1949-1960.

DeSmedt, J., and DeLey, J. (1977). Intra and intergeneric similarities of *Agrobacterium* rRNA cistrons. Int. J. Syst. Bacteriol. **27:** 222-240.

DeSmedt, J., Bauwens. M., Tytgat, R., and DeLey, J. (1980). Intra and intergeneric similarities of rRNA cistrons of free-living nitrogen-fixing bacteria. Int. J. Syst. Bacteriol. **30:** 106-122.

DeSoete, G. (1983) A least squares algorithm for fitting additive trees to proximity data. Psychometrika **48:** 621-626.

DeVos, P., and DeLey, J. (1983). Intra and intergeneric similarities of *Pseudomonas* and *Xanthomonas* rRNA cistrons. Int. J. Syst. Bacteriol. **33:** 487-509.

DeVos, P., Goor, R., Gillis, M., and DeLey, J. (1985). Ribosomal RNA cistron similarities of phytopathogenic *Pseudomonas* spp. Int. J. Syst. Bacteriol. **35:** 169-184

DeVos, P., Van Landschoot, A., Segers, P., Tytgat, R., Gillis, M., Bauwens. M., Rossau, R., Goor, R., Pot, B., Kersters, K., Lizzaraga, P., and DeLey, J. (1989). Genotypic relationships amd taxonomic localization of unclassified *Pseudomonas* and *Pseudomonas*-like strains by DNA:rRNA hybridizations. Int. J. Syst. Bacteriol. **39:** 35-49.

Dewhirst, F.E., Paster, B.J., and Bright, P.L.,(1989). 16S rRNA sequence comparison. Int. J. Syst. Bacteriol. **39:** 258-266.

Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N., and Woese, C.R. (1980). The phylogeny of prokaryotes. Science **209:** 457-463.

Gavini, F., Holmes, B., Izard, D., Beji, A., Bernigaud, A., and Jakubczak, E. (1989). Numerical taxonomy of *Pseudomonas alcaligenes, P. pseudoalcaligenes, P. mendocina, P. stutzeri,* and related bacteria. Int. J. Syst. Bacteriol. **39:** 135-144.

Gottleib, P., and Rudner, R. (1985). Restriction site polymorphism of rRNA gene sets in members of the genus *Bacillus*. Int. J. Syst. Bacteriol. **35:** 244-252.

Gueho, E., Kurtzman, C.P., and Peterson, S.W. (1990). Phylogenetic relationships among species of *Sterigmatomyces* and *Fellomyces* as determined from partial rRNA sequences. Int. J. Syst. Bacteriol. **40:** 60-65.

Gutell, R.R., and Fox, G.E. (1988). Compilation of large ribosomal subunit RNA sequences presented in a structural format. Nucl. Acids Res. **16**(seq. suppl): 175-270.

Hartmann, R.K., Toschka, H.Y., Ulbrich, N., and Erdmann, V.A. (1986). Organization of rDNA in *P. aeruginosa*.. FEBS Lett. **195:** 187-193.

Hori, H. (1975). The evolution of 5S RNA. J. Mol. Evol. **7:** 75-86.

Hori, H. and Osawa, S. (1979). Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S RNA species. Proc. Natl. Acad. Sci. USA. **76:** 381-385.

Horton, J.S., and Horgen, P.A. (1989). Polymorphisms in the nuclear DNA of *Achyla* species: some taxonomic implications. Can. J. Microbiol. **35:** 1146-1155.

Höpfl, P., Ludwig, W., Schleifer, K.H., and Larsen, N. (1989). The 23S rRNA higher-order structure of *Pseudomonas cepacia* and other prokaryotes. Eur. J. Biochem. **185:** 355-364.

Jarvis, B.D.W., Gillis, M., and DeLey, J., (1986). Intra and intergeneric similarities between the rRNA cistrons of *Rhizobium* and *Bradyrhizobium* spp. and some related bacteria. Int. J. Syst. Bacteriol. **36:** 129-138.

Johnson, J.L., and Harich, B. (1983). Comparisons of procedures for determining ribosomal RNA similarities. Curr. Microbiol. **9:** 111-120.

Johnson, J.L., and Palleroni, N.J. (1989). DNA similarities among *Pseudomonas* species. Int. J. Syst. Bacteriol. **39:** 230-235.

Keen, C.L., Mendelovitz, S., Cohen. G., Aharonowitz, Y., and Roy, K.L. (1988). Isolation and characterization of linear plasmid from *Streptomyces clavuligerus*. Mol. Gen. Genet. **212:** 172-176.

Klotz, L.C., and Blanken, R.L. (1981). A practical method for calculating evolutionary trees from sequence data. J. Theor. Biol. **91:** 261-272.

Lane, D.J., Pace, B., Olsen, G., Stahl, D.A., Sogin, M.L., and Pace, N. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc. Natl. Acad. Sci. USA. **82:** 6955-6959.

Lazo, G.R., Roffey, R., and Gabriel, D.W. (1987). Pathovars of *Xanthomonas campestris* are distinguishable by restriction fragment-length polymorphism. Int. J. Syst. Bacteriol. **37:** 214-221.

Leffers, H., Kjems, J., Østergaard, L., Larsen, N., and Garrett, R.G. (1987). Evolutionary relationships amongst Archaebacteria. A comparative study of 23S RNAs of a sulphur-dependant extreme thermophile, an extreme halophile and a thermophilic methanogen. J. Mol. Biol. **195:** 43-61.

Levy-Frebault, V.V., Thorel, M-F., Varnerot, A., and Gicquel, B. (1989). DNA polymorphism in *Mycobacterium paratuberculosis*, "Wood pigeon Mycobacteria," and related Mycobacteria analyzed by field inversion gel electrophoresis. J. Clin. Microbiol. **27:** 2823-2826.

Lim, P-O., and Seara, B.B. (1989). 16S rRNA sequence indicates that plant-pathogenic mycoplasmalike organisms are evolutionarily distinct from animal mycoplasmas. J. Bacteriol. **171:** 5901-5906.

Lindgren, P.B., Panopolous, N.J., Staskawicz, B.J., and Dahlbeck, D. (1988). Genes required for pathogenicity and hypersensitivity are conserved and interchangeable among pathovars of *Pseudomonas syringae*. Mol. Gen. Genet. **211:** 499-506.

Moureau, P., Derclaye, I., Gregoire, D., Janssen, M. and Cornelis, G.R. (1989). *Campylobacter* species identification based on polymorphism of DNA encoding rRNA. J. Clin. Micro. **27:** 1514-1517.

Noller, H.F., Kop, J.A., Wheaton, V., Brosius, J., Gutell, R.R., Kopylov, A.M., Dohme, F., and Herr, W. (1981) Secondary structure model for 23S ribosomal RNA. Nucl. Acids Res. **9:** 6167-6189.

Palleroni, N.J. (1983). The taxonomy of bacteria. BioSci. **33:** 370- 377.

Palleroni, N.J. (1984). Family I. Pseudomonadaceae. In In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 141-181. Edited by N. R. Krieg. Baltimore: Williams and Wilkins.

Palleroni, N.J., Ballard, R.W., Ralston, E., and Doudoroff, M. (1972). DNA homologies among some *Pseudomonas* spp. J. Bacteriol. **110:** 1-11.

Peladan, F., Turlot, J.C., and Monteil, H. (1984). Discriminant analysis of volatile fatty acids produced in culture medium: a novel approach to the identification of *Pseudomonas* spp. J. Gen. Microbiol. **130:** 3175-3182.

Poh, C.L., Ocampo, J.C., Sng, E.H., and Bygdeman, S.M. (1989). Rapid *in situ* generation of DNA restriction endonuclease patterns for *Neisseria gonorrhoeae*. J. Clin. Microbiol. **27:** 2784-2788.

Rigaud, G., Grange,T., and Pictet, R. (1987). The use of NaOH as tranfer solution of DNA onto nylon membrane decreases the hybridization efficiency. Nucl. Acids Res. **15:** 857.

Sands, D.C., Scroth, M.N., and Hildebrand, D.C. (1970). Taxonomy of the phytopathogenic pseudomonads. J. Bacteriol. **101:** 9-23.

Scherrer, K. and Darnell, J.E. (1962). Sedimentation characteristics of rapidly labelled RNA from HeLa cells. Biochem. Biophys. Res. Commun. **7:** 486-490.

Schleifer, K.H., Leuteritz, M., Weiss, N., Ludwig, W., Kirchhof, G., and Seidel-Rufter, H. (1990). Taxonomic study of anaerobic, gram-negative, rod-shaped bacteria from breweries: emended description of *Pectinatus cerevisiiphilus* and description of *Pectinatus frisingensis* sp. nov., *Selenomonas lacticifex* sp. nov., *Zymophilus raffinosivorans* gen. nov., sp. nov., and *Zymophilus paucivorans* sp. nov. Int. J. Syst. Bacteriol. **40:** 19-27.

Shen, C., and Maniatis, T. (1980). The organization of repetitive sequences in a cluster of rabbit β-like globin sequences. Cell **19:** 379-391.

Stackebrandt, E., and Charfreitag, O. (1990). Partial 16S rRNA primary structure of five *Actinomyces* species: phylogenetic implications and development of an *Actinomyces israelii*-specific oligonucleotide probe. J. Gen. Microbiol. **136:** 37-43.

Stahl, D.A., Lane, D.J., Olsen, G.J., and Pace, N.R. (1984). Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences. Science **224:** 409-411.

Staley, J. T. and Kreig, N. R. (1984). Classification of procaryotic organisms: an overview. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 1-5. Edited by N. R. Krieg. Baltimore: Williams and Wilkins.

Toschka, H.T., Höpfl, P., Ludwig, W., Shleifer, K.H., Ulbrich, N., and Erdmann, V.A. (1987). Complete nucleotide sequence of a 23S ribosomal RNA gene from *Pseudomonas aeruginosa* . Nucl. Acids Res. **15:** 7182.

Toschka, H.T., Höpfl, P., Ludwig, W., Shleifer, K.H., Ulbrich, N., and Erdmann, V.A., (1988). Complete nucleotide sequence of a 16S ribosomal RNA gene from *Pseudomonas aeruginosa* . Nucl. Acids Res. **16:** 2348.

Tsuji, K., Tsein, H.C., Hanson, R.S., DePalma, S.R., Scholtz, R., and LaRoche, S. (1990). 16S rRNA sequence analysis for determination of phylogenetic relationships among methylotrophs. J. Gen. Microbiol. **136:** 1-10.

Van den Eynde, H., Van de Peer, Y., Perry, J., and De Wachter, R. (1990). 5S rRNA sequences of representatives of the genera *Chlorobium*, *Prosthecochloris*, *Thermomicrobium*, *Cytophaga*, *Flavobacterium*, *Flexibacter* and *Saprospira* and a discussion of the evolution of eubacteria in general. J. Gen. Microbiol. **136:** 11-18.

Weisberg, W.G., Dobson, M.E., Samuel, J.E., Dasch, G.A., Mallavia, L.P., Baca, O., Mandelco, L., Sechrest, J.E., Weiss, E., and Woese, C.R., (1989).Phylogenetic diversity of the Rickettsiae. J. Bacteriol. **171:** 4202-06.

Willems, A., Busse, J. A., Goor, R., Pot, B., Falsen, E., Jantzen, E., Hoste, B., Gillis, M., Kersters, K., Auling, G., and DeLey, J., (1989). *Hydrogenophaga*, a new genus of hydrogen oxidizing bacteria. Int. J. Syst. Bacteriol. **39:** 319-333.

Woese, C.R. (1987). Bacterial evolution. Microbiol. Rev. **51:** 221-271.

Zhao, H., Yang, D., Woese.C., and Bryant, M.P. (1990). Assignment of *Clostridium bryantii* gen. nov., comb. nov. on the basis of a 16S rRNA sequence analysis of its crotonate-grown pure culture. Int. J. Syst. Bacteriol. **40:** 40-44.